THE ARCSINE LAW AS A UNIVERSAL AGING SCHEME FOR TRAP MODELS

GÉRARD BEN AROUS AND JIŘÍ ČERNÝ*

ABSTRACT. We give a general proof of aging for trap models using the arcsine law for stable subordinators. This proof is based on abstract conditions on the potential theory of the underlying graph and on the randomness of the trapping landscape. We apply this proof to aging for trap models on large two-dimensional tori and for trap dynamics of the Random Energy Model on a broad range of time scales.

1. INTRODUCTION

We establish in this article a general mechanism explaining the phenomenon of aging for some important dynamics in random media, i.e. the "trap models". This general scheme is based on the classical arcsine law for stable subordinators. The general context is the following. Consider $Y(i)$ a discrete-time Markov chain on a discrete space $\mathcal{V}$ with transition kernel $p(x, y)$. Assume that a function $\boldsymbol{\tau}$ is given from the state space $\mathcal{V}$ to $(0, \infty)$. $\boldsymbol{\tau}$ will later be assumed to be random. It should be seen as a random landscape or a random scenery. Consider then the following sampling process: every time the Markov chain Y is at $x \in \mathcal{V}$ it collects an (independent) exponential random variable with mean τ_x. More precisely let $(e_i : i \geq 0)$ be an independent collection of i.i.d. exponential mean-one random variables and define for any $k \geq 0$

$$S(u) = \sum_{i=0}^{\lfloor u \rfloor - 1} e_i \tau_{Y(i)}. \tag{1.1}$$

The question we address first is the following: what is the behaviour of the process $S(u)$ for large u's, when the landscape $\boldsymbol{\tau}$ is highly heterogeneous, typically (but not necessarily) when the τ_x are i.i.d. and heavy tailed? We are not interested here in the case where τ is random but reasonably tame, which is usually studied under the name of Random Walk in Random Scenery (for recent important results see [AC06, GvdHK06, GKS06] and references therein). We will, on the contrary, isolate general conditions bearing both on the distribution of the random landscape $\boldsymbol{\tau}$ and on the potential theory of the chain Y which will ensure that the process $S(u)$ can be approximated, in appropriate large time scales, by a stable subordinator. This convergence result will enable us to give a general mechanism explaining aging for the so-called Random Hopping Times dynamics or trap models, i.e. for the continuous time Markov chain $X(t)$ whose jump rates are

given by:
$$w_{xy}^{\boldsymbol{\tau}} = \tau(x)^{-1}p(x,y). \tag{1.2}$$
Indeed $X(t)$ is a time change of $Y(k)$ and the time change is given by (the inverse of) $S(u)$, which we call here the *clock process*:
$$X(t) = Y(k) \qquad \text{for all } t \in [S(k), S(k+1)). \tag{1.3}$$

We will give a limit theorem stating that the clock process is close to a stable subordinator. We will then use the classical arcsine law, which gives the probability that a deterministic interval (a, b) does not intersect the range of a subordinator as a function of the ratio a/b, in order to estimate the probability
$$R(t_w, t_w + t; \boldsymbol{\tau}) := \mathbb{P}[X(t_w + t) = X(t_w)|\boldsymbol{\tau}] \tag{1.4}$$
for large times t_w and t in appropriate time scales. In particular, we will show that it is asymptotically a function of the ratio of these times t_w/t, which is what is usually called aging for the chain $X(t)$ in the statistical physics literature.

The class of examples we do have in mind is the simple case where Y is the standard random walk on a (connected) graph $G = (\mathcal{V}, \mathcal{E})$, i.e. where the transition probability is given by:
$$p(x, y) = 1/d_x \qquad \text{if } x \text{ and } y \text{ are neighbours on the graph, } \langle x, y \rangle \in \mathcal{E} \tag{1.5}$$
and d_x is the degree of the vertex x. This class of examples comes under the name of "trap models" in statistical mechanics of disordered media, and has been introduced by J.P Bouchaud (see [Bou92]). We refer to our lecture notes [BČ06] for a complete survey and a more extensive bibliography.

Let us summarise very briefly here the state of known results about aging for trap models. The trap models have been already studied on $\mathbb{Z}^d$ and on "mean-field" objects, that is on large complete graphs. Aging was first proved for large complete graphs in [Bou92, BD95]. This case was seen by Bouchaud as a good ansatz for the dynamics of the simplest spin-glass, the Random Energy Model (REM). It corresponds to the simplest case, where the Markov chain $Y(i)$ is simply a sequence of i.i.d. random variables uniformly distributed on a large finite set. Aging was then proved for the longest possible time scales for the REM dynamics in [BBG03a, BBG03b], with a hard proof based on renewal theory. It was also proved for the trap model on $\mathbb{Z}$ with a proof based on a direct scaling limit argument in [FIN02] and [BČ05]. Finally, aging was proved for the model on $\mathbb{Z}^d$, $d \geq 2$, on the shortest possible time scale [BČM06, Čer03]. The proof there is based on a difficult coarse-graining procedure.

The striking fact is that these aging results are identical for $\mathbb{Z}^d$, $d \geq 2$ and the large complete graph, or the REM. In other terms, the mean-field results are valid from infinite dimension down to dimension 2.

The new approach we give here is based on what we have learnt from these examples and has the following advantages

1. It shows very clearly how aging is based on the arcsine law, isolating the natural interplay between the potential theory of the chain and the randomness of the landscape. It also shows that even though Bouchaud's ansatz (the model on the complete graph) is not universal in finite dimensions, one of its features is, at least for $d \geq 2$, i.e. the nature of the clock process. The two-time function R being insensitive to anything but

the range of the clock process, this is enough to imply aging, through the arcsine law, as soon as the approximation by a stable subordinator is valid.

2. It allows us to give aging results in a broad range of time scales. For instance on very long time scales in finite dimensions ($d \geq 2$), or in short time scales for the REM. We will see that there is a natural range of space scales (i.e. level sets of τ) and time scales in which the scheme based on the arcsine law applies. We will also see that it is possible in our scheme to have varying exponents in the arcsine law in varying time scales for the same model (the REM for instance). Moreover this will show that the aging phenomenon is a question about the *transient part of relaxation* to equilibrium and not necessarily related to equilibrium questions. The most striking illustration being the fact that we prove that aging can occur for the REM *above the critical temperature* (where equilibrium questions are trivial).

3. It allows us to think of random landscapes which would not be i.i.d., and thus open the possibility of studying aging of trap dynamics of more relevant spin glasses than the REM. At least with a good dose of optimism, and naturally for time scales short enough to not yet feel the model-specific equilibrium features.

Let us now be more specific, and describe precisely the questions we address here. Even though everything we prove in this paper is valid in the general context described above with no change at all, we will restrict the exposition to the case of trap models for the sake of simplicity. We will nevertheless expand a bit the framework by considering a sequence of graphs rather than one fixed graph (in order to accommodate also the large complete graphs of Bouchaud's ansatz or the hypercube in high dimensions needed for the REM). Consider thus a sequence of connected graphs $G_n = (\mathcal{V}_n, \mathcal{E}_n)$, $n \in \mathbb{N}$, with the vertex set $\mathcal{V}_n$ and the edge set $\mathcal{E}_n$. Let $(Y_n(j) : j \geq 0)$ be a discrete-time simple random walk on G_n. For each vertex $x \in \mathcal{V}_n$, let τ_x be a non-negative real number, which we call the depth of the trap at x, and denote by $\boldsymbol{\tau}_n$ the collection of these depths, $\boldsymbol{\tau}_n = \{\tau_x : x \in \mathcal{V}_n\}$. We will assume that $\boldsymbol{\tau}_n$ is a sequence of "random environments", i.e. random variables with distribution μ_n on $[0, \infty)^{\mathcal{V}_n}$. We suppose that μ_n are defined on a common probability space, so that we can consider a.s. convergence. Note that we do not assume a priori that the τ_x's are i.i.d., even though in the classical examples the τ_x's are i.i.d. and heavy tailed [BČ06].

Given the environment $\boldsymbol{\tau}_n$, we define the trap model as a continuous-time Markov process $X_n(\cdot)$ with state space $\mathcal{V}_n$ whose transition rates are given by

$$w_{xy}^{\boldsymbol{\tau}} = \begin{cases} (d_x \tau_x)^{-1} & \text{if } \langle x, y \rangle \in \mathcal{E}_n, \\ 0 & \text{otherwise.} \end{cases} \tag{1.6}$$

Here d_x stands for the degree of x in the graph G_n. In words, X_n waits at x an exponentially distributed time with mean τ_x and then it jumps to one of the neighbours of x with the equal probability d_x^{-1}. We write $\mathbb{P}_x$ for the distribution of X_n conditioned on $X_n(0) = x$. Usually, we will consider X_n to be started at some arbitrary fixed vertex that does not depend on $\boldsymbol{\tau}_n$. This vertex is denoted by $\mathbf{0}$, we write $\mathbb{P} = \mathbb{P}_{\mathbf{0}}$.

As we have already explained, X_n is a time change of Y_n. Indeed, define the *clock process* $S_n(u)$, $u \geq 0$ by

$$S_n(u) := \sum_{i=0}^{\lfloor u \rfloor - 1} e_i \tau_{Y_n(i)}. \tag{1.7}$$

So that,

$$X_n(t) = Y_n(j) \qquad \text{for all } t \in [S_n(j), S_n(j+1)). \tag{1.8}$$

To study aging we need choose a two-time function that reflects the behaviour of the system in the time interval $[t_w, t_w + t]$. The most natural two-time function for the trap models is the probability that at both times t_w and $t_w + t$ the system is in the same state,

$$R_n(t_w, t_w + t; \boldsymbol{\tau}_n) := \mathbb{P}[X_n(t_w + t) = X(t_w) | \boldsymbol{\tau}_n]. \tag{1.9}$$

There are other possible choices for the two-time function (see [BČ06]). We will however not consider them here.

Definition 1.1. We say that the function R_n exhibits aging if for some sequence $t(n)$ such that $\lim_{n \to \infty} t(n) = \infty$ it satisfies

$$\lim_{n \to \infty} R_n(t(n), (1 + \theta)t(n); \boldsymbol{\tau}_n) = R(\theta) \tag{1.10}$$

for all $\theta > 0$ and some non-trivial function $R(\theta)$. We call $R(\theta)$ the aging function.

As we have already remarked, in all cases where aging of R_n was proved, with the exception of the one-dimensional case, the limiting function $R(\theta)$ was given by the arcsine law for Lévy processes[1],

$$R(\theta) = \mathsf{Asl}_\alpha(1/1 + \theta), \tag{1.11}$$

where $\mathsf{Asl}_\alpha(u)$ stands for the distribution function of the generalised arcsine law with parameter α,

$$\mathsf{Asl}_\alpha(z) := \frac{\sin \alpha \pi}{\pi} \int_0^z u^{\alpha - 1}(1 - u)^{-\alpha} \, du. \tag{1.12}$$

Note that $\mathsf{Asl}_\alpha(a/b)$ is equal to the probability that the range of an α-stable subordinator does not intersect the interval $[a, b]$ [Ber96].

The aim of this paper is to give a set of possibly simple conditions that guarantee for general graphs G_n and time scales $t(n)$ the same behaviour, that is the convergence of $R_n(t(n), (1 + \theta)t(n))$ to $\mathsf{Asl}_\alpha(1/1 + \theta)$.

We will give first a set of four general conditions (A)–(D) which ensure the convergence of the rescaled clock process to a stable subordinator. These four conditions are true for every known example (except, naturally, for the model on $\mathbb{Z}$ where the clock process converge to a Kesten-Spitzer process [KS79], see also [BČ06]). We will then, in Section 2, give a set of four more general conditions 1–4 which are weaker but sufficient to ensure the convergence of the range of the clock process, which is enough to apply the arcsine law. In order to prove aging for the two-time function R_n using this arcsine law we still need to impose two extra technical conditions, either on top of (A)–(D) or on top of 1–4.

---

[1]In [BBG03b] a more complicated limiting procedure than in (1.10) was used and the two-time function R was slightly different, we will make more comments about this issue later.

Let us introduce some notations useful in order to state our conditions. Let T_n be a stopping time for Y_n. We use $G_{T_n}^n(x, y)$, $x, y \in \mathcal{V}_n$, to denote the Green's function of Y_n, that is the mean time that Y_n spends in y before T_n when started at x,

$$G_{T_n}^n(x, y) = \mathbb{E}_x \sum_{i=0}^{T_n-1} \mathbb{1}\{Y_n(i) = y\}. \tag{1.13}$$

For a set $A \subset \mathcal{V}_n$ we define its hitting time $H_n(A)$ as

$$H_n(A) := \inf\{i \geq 0 : Y_n(i) \in A\}. \tag{1.14}$$

To simplify the notation we write $G_A^n(\cdot, \cdot)$ for $G_{H(A)}^n(\cdot, \cdot)$. We define another two-time function $R_A^n(t_w, t_w + t)$ as the probability that X does not visits any "fresh" site in A during the observation interval: let $\ell_A(t_w)$ be the last time when X_n visited A before t_w,

$$\ell_A(t_w) = \max\{t \leq t_w : X_n(t) \in A\}, \tag{1.15}$$

and let $T = \inf\{s \geq t_w : X_n(s) \in A \setminus \{X_n(\ell_A(t_w))\}\}$, then

$$R_A^n(t_w, t_w + t; \boldsymbol{\tau}) = \mathbb{P}[T > t_w + t | \boldsymbol{\tau}]. \tag{1.16}$$

A two-point function of this type was considered in [BBG03a].

We further say that the set $A \subset \mathcal{V}_n$ is a *Poisson cloud* on $\mathcal{V}_n$ with density $\rho \in (0, 1)$ if each site $x \in \mathcal{V}_n$ is in A with probability ρ independently of all others, i.e. if A is a site-percolation process on $\mathcal{V}_n$.

We can now formulate the first set of conditions that implies aging on the time scale $t(n)$. First, we need to control the behaviour of the random environment.

Condition (A). For all n the random environment $\boldsymbol{\tau}_n$ is i.i.d. Further, there exist a depth scale $g(n)$, a density scale $\rho(n)$ and a constant $\alpha \in (0, 1)$ such that $g(n) \to \infty$ and $\rho(n) \to 0$ as $n \to \infty$, and

$$\rho(n)^{-1} \mu_n[\tau_x \geq ug(n)] \xrightarrow{n \to \infty} u^{-\alpha}, \tag{1.17}$$

uniformly on all compact subsets of $(0, \infty)$. Moreover, there exist a constant C such that for all $u > 0$ and $n \in \mathbb{N}$

$$\mu_n[\tau_x \geq ug(n)] \leq Cu^{-\alpha}\rho(n). \tag{1.18}$$

The next two conditions control the motion of the simple random walk Y_n between points of a Poisson cloud.

Condition (B). Let A_n be a sequence of Poisson clouds on $\mathcal{V}_n$ with densities $\rho\rho(n)$, $\rho \in (0, \infty)$. Then there exists a constant $\mathcal{K}_G \in (0, \infty)$ independent of ρ such that for a.e. sequence A_n

$$\max_{x \in A_n} \left| f(n)^{-1} G_{A_n \setminus \{x\}}^n(x, x) - \mathcal{K}_G \right| \xrightarrow{n \to \infty} 0, \tag{1.19}$$

where the scale $f(n)$ is given by $f(n) = t(n)/g(n)$.

Condition (C). There exists $\mathcal{K}_r \in (0, \infty)$ such that for all $s > 0$ and a.e. sequence A_n of Poisson clouds as in Condition (B)

$$\max_{x \in A_n \cup \{\mathbf{0}\}} \left| \mathbb{E}_x \left[\exp\left(-\frac{s}{r(n)} H_n(A_n \setminus \{x\}) \right) \right] - \frac{\mathcal{K}_r \rho}{s + \mathcal{K}_r \rho} \right| \xrightarrow{n \to \infty} 0, \tag{1.20}$$

where $r(n)\rho(n) = f(n)$. In other words, $H_n(A_n \setminus \{x\})/r(n)$ is asymptotically exponentially distributed with mean $1/\mathcal{K}_r\rho$. (The scale $r(n)$ represents the number of steps that X_n makes before time $t(n)$.)

Finally, we need one technical condition

Condition (D). There exists a large constant $\mathcal{K}_s$ such that for all $m > 0$ and n large

$$\sum_{x \in \mathcal{V}_n} \left(e^{\lambda_n G^n_{mr(n)}} - 1\right) \leq \mathcal{K}_s \lambda_n \sum_{x \in \mathcal{V}_n} G^n_{mr(n)}(0, x) = \mathcal{K}_s \lambda_n mr(n), \qquad (1.21)$$

and $\sum_{n=1}^{\infty} \exp\left(-c\lambda_n f(n)\right)$ is finite for all $c > 0$.

Theorem 1.2. *Assume that Conditions (A)–(D) hold. Then*

(i) for a.e. random environment $\boldsymbol{\tau}$ the rescaled clock process $t(n)^{-1}S(r(n)\cdot)$ converges to an α-stable subordinator weakly in the Skorokhod topology on $D([0, T])$ for all $T > 0$.

(ii) Further, define the set of deep traps $T^M_\varepsilon(n) := \{x \in \mathcal{V}_n : \varepsilon g(n) \leq \tau_x < Mg(n)\}$. Then $R^n_{T^M_\varepsilon(n)}$ ages:

$$\lim_{\substack{\varepsilon \to 0 \\ M \to \infty}} \lim_{n \to \infty} R^n_{T^M_\varepsilon(n)}(t(n), (1+\theta)t(n)) = \mathsf{Asl}_\alpha(1/1+\theta). \qquad (1.22)$$

2. Graph-independent mechanism of proof

In this section we state the second set of the conditions. This set is adjusted to prove aging for the two-time function R and its convergence to Asl_α. The conditions are more complicated than the Conditions (A)–(D) of the first set, and can be rather regarded as parts of a mechanism of a proof of aging.

To understand these conditions it is useful to keep in mind the analogy with the sum of i.i.d. non-negative α-stable random variables. It is known fact that, after a proper renormalisation, this sum converge to an α-stable subordinator. Moreover, the sum is typically dominated by a finite number of large contributions whose size depend on the number of terms. We want to prove that the same holds for the clock process.

To formulate the second set of conditions it is necessary to choose several objects that depend on the particular sequence G_n and on the observation time scale $t(n)$.

First, it is necessary to fix a (random) time ξ_n up to which we observe Y_n. This time will serve as an upper time scale up to which we observe Y_n. It must therefore be chosen large enough to ensure that (with high probability) $S_n(\xi_n)$ is larger than $(1+\theta)t(n)$ (see Condition 4 below). On the other hand, ξ_n should be as small as possible to simplify the verification of the other conditions.

Second, a scale $g(n)$ for deep traps should be chosen according to G_n and $t(n)$. This scale defines the set of the *deep traps* by

$$T^M_\varepsilon(n) := \{x \in \mathcal{V}_n : \varepsilon g(n) \leq \tau_x < Mg(n)\}. \qquad (2.1)$$

This set will determine the behaviour of the clock process S_n at the time scale $t(n)$. That is the clock process should be dominated by a very small number of relatively large contributions due to visits of deep traps in $T^M_\varepsilon(n)$.

We use $T_M(n) := T^\infty_M(n) = \{x \in \mathcal{V}_n : Mg(n) \leq \tau_x\}$ to denote the set of *very deep traps*. Similarly, we write $T^\varepsilon(n) := T^\varepsilon_0(n) = \{x \in \mathcal{V}_n : \tau_x < \varepsilon g(n)\}$ for the set of *shallow traps*. To justify the analogy with the sum of i.i.d. random variables, the contribution

of these two sets should be negligible. This is the content of the first two conditions. Condition 1 states that shallow traps are irrelevant because the time spent in them is too small. Condition 2 states that very deep traps are irrelevant because they are not seen by the Markov chain.

Condition 1. There is a function $h(\varepsilon)$ satisfying $\lim_{\varepsilon \to 0} h(\varepsilon) = 0$, such that for a.e. realisation of $\tau := \{\tau_n : n > 0\}$ and for all n large enough

$$\mathbb{E}\left[\sum_{i=0}^{\xi_n} e_i \tau_{Y_n(i)} \mathbb{1}\{Y_n(i) \in T^\varepsilon(n)\} \Big| \tau\right] \leq h(\varepsilon) t(n). \tag{2.2}$$

That is, the expected time spent in shallow traps before ξ_n is small with respect to $t(n)$.

The second conditions ensures the negligibility of the very deep traps. Recall that $H_n(A)$ denotes the hitting time of the set A by Y_n.

Condition 2. Given ξ_n, for any $\delta > 0$ there exists M large enough such that for a.e. realisation of τ and for all n large

$$\mathbb{P}\left[H_n(T_M(n)) \leq \xi_n \big| \tau\right] \leq \delta. \tag{2.3}$$

We need other definitions to state conditions that guarantee the existence of the limit in (1.10). First, let $r_n(j)$ be the sequence of times when a new deep trap is visited, $r_n(0) = 0$, and

$$r_n(i) = \min\left\{j > r_n(i-1) : Y_n(j) \in T_\varepsilon^M(n) \setminus \{Y_n(r_n(i-1))\}\right\}. \tag{2.4}$$

We use ζ_n to denote the largest j such that $r_n(j) \leq \xi_n$,

$$\zeta_n := \max\{j : r_n(j) \leq \xi_n\}. \tag{2.5}$$

We define the process $U_n(j)$ that records the trajectory of Y_n (and thus of X_n) restricted to the deep traps,

$$U_n(j) := Y_n(r_n(j)), \qquad j \in \mathbb{N}_0. \tag{2.6}$$

Finally, let $s_n(j)$ be the time that X_n spends at site $U_n(j)$ between steps $r_n(j)$ and $r_n(j+1)$,

$$s_n(j) := \sum_{i=r_n(j)}^{r_n(j+1)} e_i \tau_{Y_n(i)} \mathbb{1}\{Y_n(i) = U_n(j)\}, \qquad j < \zeta_n. \tag{2.7}$$

It is easy to observe that $s_n(j)$ has an exponential distribution with mean

$$\tau_{U_n(j)} G^n_{T_\varepsilon^M \setminus \{U_n(j)\}}(U_n(j), U_n(j)). \tag{2.8}$$

Since Conditions 1 and 2 ensure that the visits of the deep traps determine the behaviour of the time change $S_n(j)$, the sum $\sum_{i=1}^{j-1} s_n(i)$ can be considered as a good approximation of $S_n(r_n(j))$. We would like to show that the $s_n(j)$ become independent as $n \to \infty$, and that they have an appropriate tail behaviour. To this end, we define $(\sigma_\varepsilon^M(i), i \in \mathbb{N})$ as a sequence of i.i.d. random variables taking values between ε and M with common distribution function

$$\mathbb{P}[\sigma_\varepsilon^M(i) \leq u] = \frac{\varepsilon^{-\alpha} - u^{-\alpha}}{\varepsilon^{-\alpha} - M^{-\alpha}} =: \frac{\varepsilon^{-\alpha} - u^{-\alpha}}{p_\varepsilon^M}, \qquad u \in [\varepsilon, M]. \tag{2.9}$$

Let $(\hat{e}_i, i \in \mathbb{N})$ be a sequence of mean-one i.i.d. exponential random variables that are independent of σ_ε^M, and let $s_\infty(i) := \hat{e}_i \sigma_\varepsilon^M(i)$. For notational convenience we define $s_n(j) = s_\infty(j)$ for all $j \geq \zeta_n$.

The following conditions will ensure that the limit in (1.10) is given by the arcsine law (1.11). First, we need the asymptotic independence and the proper tail behaviour:

Condition 3. There exists a constant $\mathcal{K} > 0$ such that for all ε, M and for a.e. $\boldsymbol{\tau}$, the sequence $(s_n(j)/t(n), j \in \mathbb{N})$ converges as $n \to \infty$ in law to the sequence of i.i.d. random variables $(\mathcal{K}s_\infty(j), j \in \mathbb{N})$.

We need also to ensure that $S_n(r_n(\zeta_n))$ is larger than $(1+\theta)t(n)$ with a large probability. Since $r_n(\zeta_n) \geq \sum_{i=1}^{\zeta_n - 1} s_n(i)$, and $s_n(i)$ are easier to control than $S_n(r_n(j))$ we require

Condition 4. For a.e. $\boldsymbol{\tau}$ and for any fixed $\theta > 0$, $\delta > 0$ it is possible to choose ξ_n such that for all ε small and M large enough, and for ζ_n defined in (2.5)

$$\mathbb{P}\Big[\sum_{i=1}^{\zeta_n - 1} s_n(i) \geq (1+\theta)t(n)\Big|\boldsymbol{\tau}\Big] \geq 1 - \delta. \tag{2.10}$$

The next pair of conditions is, in principle, necessary only for a "post-processing". If they are not verified, it is possible to prove aging for the $T_\varepsilon^M(n)$-*dependent* two-time function $R_{T_\varepsilon^M(n)}^n$. Observe that this function can be also written as

$$\begin{aligned} R_{T_\varepsilon^M(n)}^n(t_w, t_w + t; \boldsymbol{\tau}) &= \mathbb{P}\big[\exists j : S_n(r_n(j)) \leq t_w < t_w + t < S_n(r_n(j+1))\big|\boldsymbol{\tau}\big] \\ &= \mathbb{P}\big[\{S_n(j) : j \in \mathbb{N}\} \cap (t_w, t_w + t] = \emptyset\big|\boldsymbol{\tau}\big]. \end{aligned} \tag{2.11}$$

To prove aging for the two-point function R we need to know that for any time t' between $S_n(r_n(j))$ and $S_n(r_n(j+1))$ the probability that $X_n(t') = U_n(j)$ is large. For a formal statement of this claim we need some more definitions. Let t'_n be a deterministic time sequence satisfying $t(n)/2 \leq t'_n \leq (1+\theta)t(n)$, and let $\delta > 0$. We define $j_n \in \mathbb{N}$ by

$$S_n(r_n(j_n)) \leq t'_n \leq S_n(r_n(j_n + 1)) - \delta t(n), \tag{2.12}$$

and $j_n = \infty$ if (2.12) is not satisfied for any integer. Let $A_n(\delta)$ be the event

$$A_n(\delta) := \{0 < j_n < \zeta_n\}. \tag{2.13}$$

Condition 5. For any δ it is possible to choose ε small and M large enough such that for a.e. $\boldsymbol{\tau}$ and all n large enough

$$\mathbb{P}[X_n(t'_n) = U_n(j_n)|A_n(\delta), \boldsymbol{\tau}] \geq 1 - \delta. \tag{2.14}$$

The last condition that we need to prove aging for R excludes repetitions in the sequence U_n.

Condition 6. For any fixed ε and M and a.e. $\boldsymbol{\tau}$

$$\lim_{n \to \infty} \mathbb{P}[\exists 0 < i, j \leq \zeta_n \text{ such that } i \neq j \text{ and } U_n(i) = U_n(j)|\boldsymbol{\tau}] = 0. \tag{2.15}$$

We now show how to use these six conditions to prove the aging behaviour for the two-time functions R_n, $R_{T_\varepsilon^M(n)}^n$.

Theorem 2.1. *(i) Assume that Conditions 1–6 are satisfied. Then for a.e. realisation of the random environment $\boldsymbol{\tau}$*

$$\lim_{n\to\infty} R_n(t(n), (1+\theta)t(n); \boldsymbol{\tau}) = \mathsf{Asl}_\alpha(1/1+\theta). \tag{2.16}$$

(ii) If only Conditions 1–4 hold, then the same is valid for the two-time function $R^n_{T^M_\varepsilon(n)}$,

$$\lim_{\substack{\varepsilon\to 0 \\ M\to\infty}} \lim_{n\to\infty} R^n_{T^M_\varepsilon(n)}(t(n), (1+\theta)t(n); \boldsymbol{\tau}) = \mathsf{Asl}_\alpha(1/1+\theta). \tag{2.17}$$

Before proving Theorems 1.2 and 2.1 let us explain how both sets of conditions are related.

Proposition 2.2. *Conditions (A)–(D) imply Conditions 1–4 for the same scale $g(n)$ and for $\xi_n = mr(n)$ with some large θ-dependent constant m.*

We will use this proposition for both examples that we study later. That is to prove aging for R_n we will always verify Conditions (A)–(D), 5 and 6.

Observe also that neither in Conditions 1–6, neither in Theorem 2.1 we suppose that the τ_x's are i.i.d. This assumption is however used twice when proving Proposition 2.2 as we will see later. First, we will use the independence to verify Condition 1 from Condition (D) (see formulas (2.48), (2.49) below), that is to prove that the time spent in the shallow traps is small. We do however believe that Condition 1 stays valid also for some dependent random environment. The second use of the independence is more substantial. It implies that the geometrical structure of the set of the deep traps is particularly simple: it is a Poisson cloud. It is therefore easy to control, e.g., the minimal distance between deep traps or the relative size of the slices T^u_ε, T^M_ε. This control can be problematic when the τ_x's are dependent.

Remark also that Theorem 1.2(ii) is a simple consequence of Proposition 2.2 and Theorem 2.1(b). Therefore, we first prove Theorem 2.1, then we verify Proposition 2.2. In the end we show Theorem 1.2(i), that is the convergence of the clock process. Note that Conditions 1–6 are not strong enough to imply directly such a convergence. We are however not aware of any particular case where Conditions 1–6 hold and this convergence does not take place.

Proof of Theorem 2.1. Let us define

$$\tilde{S}(j) = \frac{1}{t(n)} \sum_{i=1}^{j-1} s_n(i) \tag{2.18}$$

and let $E = E(n)$ be the event whose probability we are trying to estimate,

$$E(n) = \{X_n(t(n)) = X_n((1+\theta)t(n))\}. \tag{2.19}$$

We first explain the strategy of the proof. The most important observation is that the clock process S_n contains enough information about X_n to prove aging. Between times $S_n(r_n(j))$ and $S_n(r_n(j+1))$ the process X_n visits (possibly many times) only one deep trap, $U_n(j)$, and it also visits many shallow traps. Condition 5 ensures that if we pick a time t between $S_n(r_n(j))$ and $S_n(r_n(j+1))$, then $X_n(t) = U_n(j)$ with a high probability. That means that if $S_n(r_n(j)) \le t(n) \le (1+\theta)t(n) < S_n(r_n(j+1))$ for some $j \ge 1$, then E holds with a probability close to 1. On the other hand, if there is a j such

that $t(n) < S_n(r_n(j)) < (1+\theta)t(n)$, then, using Condition 6, E can happen only if at both times $t(n)$ and $(1+\theta)t(n)$ the process is in the same shallow trap. However, this event has (again by Condition 5) a very small probability. Therefore, it is important to estimate the probability that there is no $S_n(r_n(j))$ in the time interval of interest. To this end, we will show that $\tilde{S}_n(j)$ is a good approximation of $S_n(r_n(j))/t(n)$, and then we will estimate the probability that there is no $\tilde{S}_n(j)$ in $[1, 1+\theta]$. At the end of the proof we use these results to give a rigorous version of the reasoning in this paragraph.

First, let us show that $\tilde{S}_n(j)$ approximates well $S_n(r_n(j))/t(n)$, at least for all relevant indices $j \leq \zeta_n$.

Lemma 2.3. *For all θ and $\delta > 0$ there exist ξ_n, ε, and M such that $\boldsymbol{\tau}$-a.s.*

$$\mathbb{P}\left[\max\left\{\left|\frac{S_n(r_n(j))}{t(n)} - \tilde{S}_n(j)\right| : S_n(r_n(j)) \leq (1+\theta)t(n)\right\} > \delta \,\middle|\, \boldsymbol{\tau}\right] < \delta. \tag{2.20}$$

Proof. By Condition 4 we can choose ξ_n not depending on ε and M such that

$$\mathbb{P}[\tilde{S}_n(\zeta_n) \leq 1 + \theta] \leq \delta/2. \tag{2.21}$$

Observing that $S_n(r_n(j))/t(n) - \tilde{S}_n(j)$ is positive and increasing in j, it is sufficient to estimate $S_n(r_n(\zeta_n))/t(n) - \tilde{S}_n(\zeta_n)$. However,

$$\frac{S_n(r_n(\zeta_n))}{t(n)} - \tilde{S}_n(\zeta_n) \leq \sum_{i=0}^{\xi_n} \frac{e_i \tau_{Y_n(i)}}{t(n)} \mathbb{1}\{Y_n(i) \in T^\varepsilon(n) \cup T_M(n)\}. \tag{2.22}$$

The contribution coming from $T^\varepsilon(n)$ can be bounded using Condition 1 and Chebyshev inequality,

$$\mathbb{P}\left[t(n)^{-1} \sum_{i=0}^{\xi_n} e_i \tau_{Y_n(i)} \mathbb{1}\{Y_n(i) \in T^\varepsilon(n)\} \geq \delta/2 \,\middle|\, \boldsymbol{\tau}\right] \leq c\delta^{-1} h(\varepsilon). \tag{2.23}$$

Since $h(\varepsilon) \to 0$ as $\varepsilon \to 0$, we can fix ε such that the last expression is bounded by $\delta/4$. Similarly, using Condition 2, we can choose M such that

$$\mathbb{P}\left[\sum_{i=0}^{\xi_n} e_i \tau_{Y_n(i)} \mathbb{1}\{Y_n(i) \in T_M(n)\} \neq 0 \,\middle|\, \boldsymbol{\tau}\right] \leq \delta/4. \tag{2.24}$$

The lemma then follows combining (2.21), (2.23) and (2.24). $\qquad\square$

We further compute the probability that an interval does not contain any of the $\tilde{S}_n(j)$'s.

Lemma 2.4. *For all $0 < a < b$ and for a.e. $\boldsymbol{\tau}$*

$$\lim_{\substack{\varepsilon \to 0 \\ M \to \infty}} \lim_{n \to \infty} \mathbb{P}\left[[a, b] \cap \{\tilde{S}_n(j), j \in \mathbb{N}\} = \emptyset \,\middle|\, \boldsymbol{\tau}\right] = \mathsf{Asl}_\alpha(a/b). \tag{2.25}$$

Proof. Let $y_0 = 0$ and let $(y_i, i \in \mathbb{N})$, $y_i < y_{i+1}$, be a homogeneous Poisson point process on $(0, \infty)$ with intensity $p_\varepsilon^M := \varepsilon^{-\alpha} - M^{-\alpha}$. Consider process $(\mathcal{Y}_n(u), u \geq 0)$ given by

$$\mathcal{Y}_n(u) = \sum_{i: y_i \leq u} \frac{s_n(j)}{t(n)} \tag{2.26}$$

or, equivalently, $\mathcal{Y}_n(u) = \tilde{S}_n(j)$ for all $u \in [y_{j-1}, y_j)$. We use $\mathcal{Y}_\varepsilon^M$ to denote the Lévy process whose Lévy measure ν_ε^M is given by (a multiple of) the distribution of $s_\infty(j)$,

$$\mathbb{E}\big[\exp\big(-\lambda \mathcal{Y}_\varepsilon^M(u)\big)\big] = \exp\Big\{-u \int_0^\infty (1 - e^{-\lambda v})\nu_\varepsilon^M(\mathrm{d}v)\Big\}, \qquad (2.27)$$

$$\nu_\varepsilon^M(\mathrm{d}v) = p_\varepsilon^M \mathbb{P}[s_\infty \in \mathrm{d}v] = \mathrm{d}v \int_\varepsilon^M \alpha z^{-\alpha-2} e^{-v/z}\,\mathrm{d}z. \qquad (2.28)$$

Condition 3 implies that as $\mathcal{Y}_n \to \mathcal{K}\mathcal{Y}_\varepsilon^M$ as $n \to \infty$ weakly in the Skorokhod topology for a.e. $\boldsymbol{\tau}$.

Let $\mathcal{R}(\mathcal{Y}_n)$ denote the range of the process $\mathcal{Y}_n$, $\mathcal{R}(\mathcal{Y}_n) = \bigcup_{u \geq 0} \mathcal{Y}_n(u)$. It follows that

$$\mathbb{P}\big[[a, b] \cap \{\tilde{S}_n(j), j \in \mathbb{N}\} = \emptyset\big|\boldsymbol{\tau}\big] = \mathbb{P}\big[\mathcal{R}(\mathcal{Y}_n) \cap [a, b] = \emptyset\big|\boldsymbol{\tau}\big], \qquad (2.29)$$

Since the distribution of s_∞ has no atoms, the probability that $a/\mathcal{K}$ or $b/\mathcal{K}$ are contained in $\mathcal{R}(\mathcal{Y}_\varepsilon^M)$ is zero. Therefore, the weak convergence of $\mathcal{Y}_n$ implies that

$$\lim_{n \to \infty} \mathbb{P}\big[[a, b] \cap \{\tilde{S}_n(j), j \in \mathbb{N}\} = \emptyset\big|\boldsymbol{\tau}\big] = \mathbb{P}[\mathcal{R}(\mathcal{Y}_\varepsilon^M) \cap [a/\mathcal{K}, b/\mathcal{K}] = \emptyset|\boldsymbol{\tau}]. \qquad (2.30)$$

To estimate the right-hand side of the previous expression observe that as $\varepsilon \to 0$ and $M \to \infty$ the Lévy measure ν_ε^M converges to

$$\mathrm{d}v \int_0^\infty \alpha z^{-\alpha-2} e^{-v/z}\,\mathrm{d}z = \alpha\Gamma(1+\alpha)v^{-1-\alpha}\mathrm{d}v. \qquad (2.31)$$

This is the Lévy measure of an α-stable subordinator. Therefore, as $\varepsilon \to 0$ and $M \to \infty$, the process $\mathcal{Y}_\varepsilon^M$ converges weakly in the Skorokhod topology to the stable subordinator. Using the same reasoning as before we get

$$\lim_{\substack{\varepsilon \to 0 \\ M \to \infty}} \mathbb{P}[\mathcal{R}(\mathcal{Y}_\varepsilon^M) \cap [a/\mathcal{K}, b/\mathcal{K}] = \emptyset|\boldsymbol{\tau}] = \mathsf{Asl}_\alpha(a/b). \qquad (2.32)$$

This finishes the proof. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

We can now finally estimate the probability of the event $E(n)$ (defined by (2.19)) for large n. Fix $\delta > 0$. Let $\mathcal{B}$ be the event that is considered in Condition 4,

$$\mathcal{B} = \Big\{\sum_{i=1}^{\zeta_n} s_n(i) \geq (1+\theta)t(n)\Big\}. \qquad (2.33)$$

According to Condition 4 we can choose ξ_n such that $\mathbb{P}[\mathcal{B}^c|\boldsymbol{\tau}] \leq \delta$ for a.e. $\boldsymbol{\tau}$. We divide the probability space into three disjoint sets:

$$\begin{aligned} G_1(n) &= \{\mathrm{dist}(1, \mathcal{R}(\mathcal{Y}_n)) \leq 2\delta \text{ or } \mathrm{dist}(1+\theta, \mathcal{R}(\mathcal{Y}_n)) \leq 2\delta\} \\ G_2(n) &= \{\mathrm{dist}(1, \mathcal{R}(\mathcal{Y}_n)) > 2\delta, \mathrm{dist}(1+\theta, \mathcal{R}(\mathcal{Y}_n)) > 2\delta \text{ and} \\ &\quad (1, 1+\theta) \cap \mathcal{R}(\mathcal{Y}_n) \neq \emptyset\} \\ G_3(n) &= \{[1-2\delta, 1+\theta+2\delta] \cap \mathcal{R}(\mathcal{Y}_n) = \emptyset\} \end{aligned} \qquad (2.34)$$

Here, for $A, B \subset \mathbb{R}$, $\mathrm{dist}(A, B) = \min\{|x - y| : x \in A, y \in B\}$. As we have already remarked, this division has the following reasons. Heuristically, on the event $\mathcal{B}$, to the precision δ (see Lemma 2.3), any interval that does not intersect $\mathcal{R}(\mathcal{Y}_n)$ corresponds to a time period spent in one site of the top and neighbouring shallow traps. On the

other hand, the points of $\mathcal{R}(\mathcal{Y}_n)$ (or more precisely very short periods preceding them) correspond to times when no deep trap is visited for a large number of steps.

We wish to show that the events $E(n)$ and $G_3(n)$ are essentially the same. Obviously, for a.e. τ (omitting the conditioning on τ in the notation)

$$\mathbb{P}[E(n) \cap G_3(n)] \leq \mathbb{P}[E(n)] \leq \mathbb{P}[G_3(n)] + \mathbb{P}[G_1(n)] + \mathbb{P}[E(n) \cap G_2(n)]. \tag{2.35}$$

We should therefore estimate all quantities in the last display. The probability of $G_1(n)$ is small. Indeed,

$$\mathbb{P}[G_1(n)] \leq \mathbb{P}[\mathrm{dist}(1, \mathcal{R}(\mathcal{Y}_n)) \leq 2\delta] + \mathbb{P}[\mathrm{dist}(1 + \theta, \mathcal{R}(\mathcal{Y}_n)) \leq 2\delta]. \tag{2.36}$$

The both probabilities on the right-hand side can be estimated using Lemma 2.4, namely it is possible to choose ε small and M, n large enough, such that

$$\mathbb{P}[\mathrm{dist}(1, \mathcal{R}(\mathcal{Y}_n)) \leq 2\delta] = \mathbb{P}\big[[1 - 2\delta, 1 + 2\delta] \cap \mathcal{R}(\mathcal{Y}_n) = \emptyset\big]$$
$$\leq \delta + 1 - \mathsf{Asl}_\alpha \left(\frac{1 - 2\delta}{1 + 2\delta} \right) \leq C\delta^{1-\alpha}. \tag{2.37}$$

In the same way we estimate the second probability from (2.36).

If $\mathcal{B} \cap G_2(n)$ holds, then there are $j_1 < j_2 \leq \zeta_n$ such that

$$\tilde{S}_n(j_1) + 2\delta \leq 1 \leq \tilde{S}_n(j_1 + 1) - 2\delta,$$
$$\tilde{S}_n(j_2) + 2\delta \leq 1 + \theta \leq \tilde{S}_n(j_2 + 1) - 2\delta, \tag{2.38}$$

and therefore, using Lemma 2.3,

$$S_n(r_n(j_1)) + \delta t(n) \leq t(n) \leq S_n(r_n(j_1 + 1)) - \delta t(n),$$
$$S_n(r_n(j_2)) + \delta t(n) \leq (1 + \theta)t(n) \leq S_n(r_n(j_2 + 1)) - \delta t(n). \tag{2.39}$$

Hence, according to Condition 5, for all ε small and M large

$$\mathbb{P}\big[X(t(n)) = U_n(j_1), X_n((1 + \theta)t(n)) = U_n(j_2)\big|\tau, \mathcal{B} \cap G_2\big] \geq 1 - \delta. \tag{2.40}$$

Using Condition 6, it follows that $\mathbb{P}[E(n)|G_2(n) \cap \mathcal{B}] \leq \delta$ for all n large enough. Therefore,

$$\mathbb{P}[E(n) \cap G_2(n)] \leq \mathbb{P}[E(n)|G_2(n) \cap \mathcal{B}] + \mathbb{P}[\mathcal{B}^c] \leq 2\delta. \tag{2.41}$$

At last, we estimate the terms related to $G_3(n)$. Using Lemma 2.4, choosing ε small and n, M large enough, we get

$$\big|\mathbb{P}[G_3(n)] - \mathsf{Asl}_\alpha(1/1 + \theta)\big| \leq C\delta. \tag{2.42}$$

Finally, by Condition 5 and using a similar reasoning as in (2.38), (2.39),

$$\mathbb{P}\big[X(t(n)) = X((1 + \theta)t(n))\big|\mathcal{B} \cap G_3\big] \geq 1 - \delta. \tag{2.43}$$

Since δ can be taken arbitrarily small, Theorem 2.1(i) follows from (2.35) and the results of the last three paragraphs.

The proof of Theorem 2.1(ii) proceeds along the same lines as the proof of (i). The only needed changes are the re-definition of $E(n)$ as

$$E(n) := \big\{\exists j : S_n(r_n(j)) \leq t(n) < (1 + \theta)t(n) \leq S_n(r_n(j + 1))\big\} \tag{2.44}$$

and the observation that Conditions 5 and 6 are not necessary in this case. $\qquad\square$

Proof of Proposition 2.2. Now we show that Conditions (A)–(D) imply Conditions 1–4. To this end we choose $\xi_n = mr(n)$ with some large constant m. The scale $g(n)$ is, of course, the same in both sets of conditions.

(A), (D) $\implies$ *Condition 1.* We show that there is a large constant K independent of ε, M, and m such that Condition 1 is satisfied with $h(\varepsilon) = Km\varepsilon^{1-\alpha}$, that is for $\mathbb{P}$-a.e. realisation of $\boldsymbol{\tau}$ and for all large n

$$\mathbb{E}\Big[\sum_{i=0}^{\xi_n} e_i \tau_{Y_n(i)} \mathbb{1}\{Y_n(i) \in T^\varepsilon(n)\} \Big| \boldsymbol{\tau} \Big] \leq Km\varepsilon^{1-\alpha} t(n). \tag{2.45}$$

To this end we use the same slicing strategy as in [BČM06, Čer03]. We divide $T^\varepsilon(n)$ into disjoint sets $T_{\varepsilon 2^{-i}}^{\varepsilon 2^{-i+1}}(n)$ with $i \in \mathbb{N}$. We show that there is a large constant K' such that $\mathbb{P}$-a.s. for all but a finite number of n the following holds: for all $i \in \mathbb{N}$

$$\mathbb{E}\Big[\sum_{i=0}^{\xi_n} e_i \tau_{Y_n(i)} \mathbb{1}\{Y_n(i) \in T_{\varepsilon 2^{-i}}^{\varepsilon 2^{-i+1}}\} \Big| \boldsymbol{\tau} \Big] \leq K'm\varepsilon^{1-\alpha} 2^{i(\alpha-1)} t(n). \tag{2.46}$$

The summation over all i then yields directly the claim (2.45).

Let $p_{n,i} = \mathbb{P}[x \in T_{\varepsilon 2^{-i}}^{\varepsilon 2^{-i+1}}]$. By Condition (A)

$$p_{n,i} \leq C\varepsilon^{-\alpha} 2^{i\alpha} \rho(n) \tag{2.47}$$

for some C independent of n, i, and ε. For all $i \in \mathbb{N}$ we have

$$\mathbb{P}\Big[\mathbb{E}\Big[\sum_{i=0}^{\xi_n} e_i \tau_{Y_n(i)} \mathbb{1}\{Y_n(i) \in T_{\varepsilon 2^{-i}}^{\varepsilon 2^{-i+1}}\} \Big| \boldsymbol{\tau} \Big] \geq K'm\varepsilon^{1-\alpha} 2^{i(\alpha-1)} t(n) \Big]$$

$$= \mathbb{P}\Big[\sum_{x \in \mathcal{V}_n} G_{\xi_n}^n(\mathbf{0}, x) \tau_x \mathbb{1}\{x \in T_{\varepsilon 2^{-i}}^{\varepsilon 2^{-i+1}}\} \geq K'm\varepsilon^{1-\alpha} 2^{i(\alpha-1)} t(n) \Big] \tag{2.48}$$

$$\leq \mathbb{P}\Big[\sum_{x \in \mathcal{V}_n} G_{mr(n)}^n(\mathbf{0}, x) \mathbb{1}\{x \in T_{\varepsilon 2^{-i}}^{\varepsilon 2^{-i+1}}\} \geq K'm\varepsilon^{-\alpha} 2^{i\alpha-1} t(n) g(n)^{-1} \Big].$$

By Chebyshev inequality with λ_n of Condition (D) this is bounded by

$$\leq \exp(-\lambda_n m K' \varepsilon^{-\alpha} 2^{i\alpha-1} t(n) g(n)^{-1}) \prod_{x \in \mathcal{V}_n} \big[1 + p_{n,i}\big(e^{\lambda_n G_{mr(n)}^n(\mathbf{0},x)} - 1\big) \big]. \tag{2.49}$$

Using $x \geq \log(1+x)$ and (2.47) we get

$$\leq \exp\Big\{ -\lambda_n m K' \varepsilon^{-\alpha} 2^{i\alpha-1} t(n) g(n)^{-1} + C\varepsilon^{-\alpha} 2^{i\alpha} \rho(n) \sum_{x \in \mathcal{V}_n} \big(e^{\lambda_n G_{mr(n)}^n(\mathbf{0},x)} - 1\big) \Big\}. \tag{2.50}$$

Condition (D) ensures that the last expression is bounded by

$$\leq \exp\{ -\lambda_n m \varepsilon^{-\alpha} 2^{i\alpha} t(n) g(n)^{-1} (K' - C\mathcal{K}_s) \} \tag{2.51}$$

Now it is easy to prove (2.46) and thus Condition 1: it is sufficient to take $K' > 2C\mathcal{K}_s$, to sum over all i and then to apply the Borel-Cantelli lemma. This is possible due to the second part of Condition (D).

(A), (C) $\implies$ *Condition 2.* By the assumptions of Theorem 1.2 τ_x are i.i.d. Therefore, $T_M(n)$ is a Poisson cloud with density that is bounded by $C\rho(n)M^{-\alpha}$. The normalised

hitting time of this cloud, $H_n(T_M(n))/r(n)$, is by Condition (C) asymptotically expo-
nentially distributed with mean $M^\alpha/\mathcal{K}_r C$. So that, by choosing M large, we can make
the probability that $H_n(T_M(n)) \le \xi_n$ arbitrarily small.

(A)–(C) $\Longrightarrow$ *Condition 3.* We first prove the following easy lemma which is a conse-
quence of Condition (C) and the lack-of-memory property of the exponential distribution

Lemma 2.5. *Assume that Condition (C) holds. Let $x \in A_n \cup \{0\}$, and let A_n^1, A_n^2 be
such that $A_n^1 \cup A_n^2 = A_n$, $A_n^1 \cap A_n^2 = \emptyset$ and*

$$\lim_{n \to \infty} |A_n^k|/|A_n| = \rho_k/\rho \qquad k \in \{1, 2\}, \rho_k \in (0, 1). \tag{2.52}$$

*Define $H_k := H_n(A_n^k \setminus \{x\})/r(n)$, $H := H(A_n \setminus \{x\})/r(n)$. Then, given that $Y_n(0) = x$,
the distribution function of $H\mathbb{1}\{H_1 < H_2\}$ converges uniformly in x to $F(u) := \rho_2/\rho +
\rho_1(1 - e^{-\mathcal{K}_r \rho u})/\rho$, $u \ge 0$, and, in particular,*

$$\lim_{n \to \infty} \sup_{x \in A_n} \left| \mathbb{P}_x[H_1 < H_2] - \frac{\rho_1}{\rho} \right| = 0. \tag{2.53}$$

Proof. We define $\nu_n^k(\mathrm{d}u) := \mathbb{P}_x[H_k \in \mathrm{d}u]$, $k \in \{\emptyset, 1, 2\}$, and

$$f_n(u) = \begin{cases} \mathbb{P}_x[H_1 < H_2 | H = u] & \text{if } u \in r(n)^{-1}\mathbb{Z}, \\ 0 & \text{otherwise.} \end{cases} \tag{2.54}$$

Using a decomposition on H and $Y_n(H)$, we get from the Markov property of Y_n for any
continuous bounded function $h : \mathbb{R}^+ \to \mathbb{R}$

$$\int h(u)\nu_n^1(\mathrm{d}u) = \int \nu_n(\mathrm{d}u)f_n(u)h(u)$$
$$+ \int \nu_n(\mathrm{d}u)(1 - f_n(u)) \int \bar{\nu}_{n,u}(\mathrm{d}v)h(u + v), \tag{2.55}$$

where the measure $\bar{\nu}_{n,u}$ is defined as

$$\bar{\nu}_{n,u}(\mathrm{d}v) = \sum_{y \in A_n^2 \setminus \{x\}} \mathbb{P}_x[Y_n(r(n)H) = y | H = u, H_2 < H_1] \mathbb{P}_y[H_1 \in \mathrm{d}v]. \tag{2.56}$$

Now we take $h(u) = e^{-\eta u}$, $\eta \ge 0$. From Condition (C) it follows that for all $\varepsilon > 0$ there
is n_0 independent of $x \in A_n$ such that for $n > n_0$, $k \in \{\emptyset, 1, 2\}$,

$$\left| \int \nu_n^k(\mathrm{d}u)e^{-\eta u} - \frac{\mathcal{K}_r \rho_k}{\mathcal{K}_r \rho_k + \eta} \right| \le \varepsilon,$$
$$\left| \int \bar{\nu}_{n,u}(\mathrm{d}u)e^{-\eta u} - \frac{\mathcal{K}_r \rho_1}{\mathcal{K}_r \rho_1 + \eta} \right| \le \varepsilon. \tag{2.57}$$

Inserting this in (2.55) and re-arranging it slightly we get

$$\int e^{-\eta u} f_n(u)\nu_n(\mathrm{d}u) = \frac{\mathcal{K}_r \rho}{\mathcal{K}_r \rho + \eta} \cdot \frac{\rho_1}{\rho} + O(\varepsilon), \tag{2.58}$$

from which the first claim of the lemma directly follows. Taking $\eta = 0$ in (2.58) we get
also (2.53). $\qquad \square$

To prove Condition 3 we need to verify that the sequence $s_n(j)/t(n)$ (see (2.7) for the definition) converges to the i.i.d. sequence s_∞. We control first the sequence of depths of visited deep traps:

Lemma 2.6. *The sequence $(\tau_{U_n(j)}/g(n), j \in \mathbb{N})$ converges weakly as $n \to \infty$ to the i.i.d. sequence $(\sigma_\varepsilon^M(j), j \in \mathbb{N})$ defined in (2.9).*

Proof. Fix $u \in (\varepsilon, M)$. Since τ_x are i.i.d. the set T_ε^M and also its subsets T_ε^u, T_u^M are Poisson clouds with densities converging to $p_\varepsilon^M \rho(n)$, resp. $p_\varepsilon^u \rho(n)$ and $p_u^M \rho(n)$, where $p_a^b := a^{-\alpha} - b^{-\alpha}$. We can therefore use Condition (C) and Lemma 2.5: uniformly in $x \in T_\varepsilon^M \cup \{\mathbf{0}\}$

$$\mathbb{P}_x\left[\tau_{Y_n(H(T_\varepsilon^M \setminus \{x\}))} \leq ug(n)\right] = \mathbb{P}_x\left[H(T_\varepsilon^u \setminus \{x\}) = H(T_\varepsilon^M \setminus \{x\})\right] \xrightarrow{n \to \infty} \frac{p_\varepsilon^u}{p_\varepsilon^M}. \qquad (2.59)$$

The lemma then follows easily. $\qquad \square$

Observe now that given $U_n(j) = x \in T_\varepsilon^M(n)$, the random variable $s_n(j)/\tau_x$ is exponentially distributed with mean $G_{T_\varepsilon^M \setminus \{x\}}^n(x, x) = \mathcal{K}_G f(n)(1 + o(1))$ as $n \to \infty$. The error is uniformly bounded in x. Using the Markov property we get

$$\mathbb{P}[s_n(j)/t(n) \geq u | s_n(1), \ldots, s_n(j-1)]$$

$$= \int_\varepsilon^M \mathbb{P}[s_n(j) \geq ut(n) | \tau_{U_n(j)} = vg(n), s_n(1), \ldots, s_n(j-1)]$$

$$\times \mathbb{P}[\tau_{U_n(j)}/g(n) \in dv | s_n(i), i = 1, \ldots, j-1] \qquad (2.60)$$

$$= \int_\varepsilon^M \exp\left\{-\frac{ut(n)(1 + o(1))}{vg(n)\mathcal{K}_G f(n)}\right\} \mathbb{P}[\tau_{U_n(j)}/g(n) \in dv | s_n(i), i = 1, \ldots, j-1].$$

By definition $g(n)f(n) = t(n)$. The weak convergence of the sequence $\tau_{U_n(j)}/g(n)$ proved in Lemma 2.6 then yields that the right-hand side of (2.60) converges to

$$\int_\varepsilon^M e^{-u/\mathcal{K}_G v} \mathbb{P}[\sigma_\varepsilon^M(1) \in dv], \qquad (2.61)$$

and Condition 3 is proved.

(A)–(C) $\implies$ *Condition 4.* Fix temporarily $\varepsilon = 1/2$, $M = 2$. Since the distribution of $s_n(i)/t(n)$ converges to the distribution of $s_\infty(i)$, for any $\delta > 0$ it is possible to fix a large integer R such that for all n large enough

$$\mathbb{P}\left[\sum_{i=1}^R s_n(i) \geq (1 + \theta)t(n)\right] \geq 1 - \delta/2. \qquad (2.62)$$

To satisfy Condition 4, the constant m should be fixed such that $\mathbb{P}[\zeta_n \geq R] \geq 1 - \delta/2$. But it can be done easily, since ζ_n is the number of deep traps that are visited before ξ_n, and the distribution of this number converges to the Poisson distribution with mean $p_{1/2}^2 m/\mathcal{K}$ as follows from Condition (C). Taking now $\varepsilon < 1/2$ or $M > 2$, the sum $\sum_{i=1}^{\zeta_n} s_n(i)$ can become only larger. Therefore, for chosen m Condition 4 is verified for all $\varepsilon < 1/2$ and $M > 2$. $\qquad \square$

Proof of Theorem 1.2(i). We want to to verify that Conditions (A)–(D) imply the weak convergence in the Skorokhod topology of the process $S(r(n)\cdot)/t(n)$ to an α-stable subordinator. As usual, it is sufficient to check the convergence of fixed-time distributions and the tightness in $D([0, T])$.

Fix $t_0 \in (0, T)$. Let $N_n(t_0)$ be the number of deep traps visited in first $t_0 r(n)$ steps,

$$N_n(t_0) = \max\{j : r_n(j) \leq t_0 r(n)\}. \tag{2.63}$$

It follows from Condition (C) that the distribution of $N_n(t_0)$ converges to the Poisson distribution with mean $t_0 \mathcal{K}_r p_\varepsilon^M$. Due Proposition 2.2 Conditions 1–4 holds, therefore using a similar reasoning as to show Lemma 2.3, we can show that for any δ there are ε and M such that with probability larger than $1 - \delta$

$$t(n)^{-1} \sum_{i=0}^{N_n(t_0)-1} s_n(i) \leq t(n)^{-1} S_n(t_0 r(n)) \leq \delta + t(n)^{-1} \sum_{i=0}^{N_n(t_0)} s_n(i). \tag{2.64}$$

Now, it is easy to proceed as in Lemma 2.4 to see that the Laplace transform of $(\mathcal{K}_G t(n))^{-1} \sum_{i=0}^{N_n(t_0)} s_n(j)$ converges as $n \to \infty$ to (2.27) evaluated at $u = \mathcal{K}_r t_0$. The distribution with such Laplace transform can be made arbitrarily close (in the weak sense) to the distribution at $\mathcal{K}_r t_0$ of the α-stable subordinator with Lévy measure (2.31) by taking ε small and M large. From this the convergence of the distribution at the fixed time t_0 follows. One can get completely analogously the convergence of joint distributions of $(t(n)^{-1} S_n(r_n t_0), \ldots, t(n)^{-1} S_n(r_n t_k))$.

Since S_n is increasing to prove the tightness it is sufficient to check the tightness of the sequence of real random variables $t(n)^{-1} S_n(r(n)T)$. However, this can be done easily using (2.64) with $t_0 = T$ and Conditions 1–4. $\qquad\square$

3. AGING FOR SHORT TIME SCALES IN THE RANDOM ENERGY MODEL

The Random Energy Model (REM) is the simplest mean-field model for spin-glasses and its static behaviour is well understood. The studies of dynamics are much more sparse. The first proof of aging in the REM was given in [BBG03a, BBG03b], based on renewal theory. The approach of Section 2 allows to prove aging on a broad range of shorter time scales. We will compare both results later. Before doing it, let us define the model and give some motivation why and in what ranges of times and temperatures aging occurs.

The REM is a mean-field model of a spin-glass. It consists of n spins that can take values -1 or 1, that is configurations of the REM are elements of $\mathcal{V}_n = \{-1, 1\}^n$. The energies $\{E_x, x \in \mathcal{V}_n\}$ of the configurations are i.i.d. random variables. The standard choice of the marginal distribution of E_x is centred normal distribution with variance n. We then define

$$\tau_x = \exp(\beta E_x). \tag{3.1}$$

For the dynamics of the REM we require that only one spin can be flipped at a given moment. This corresponds to

$$\mathcal{E}_n = \{\langle x, y \rangle \in \mathcal{V}_n^2 : \sum_{i=1}^n |x_i - y_i| = 2\}, \tag{3.2}$$

where $(x_1, \ldots, x_n)$ are the values of individual spins. We use G_n to denote the n-dimensional hypercube $(\mathcal{V}_n, \mathcal{E}_n)$. There are many choices for the dynamics of REM, which have the Gibbs measure τ as a reversible measure. We will naturally consider the trap model dynamics (1.6). which is one of the simplest choices. We always suppose that

$$Y_n(0) = X_n(0) = \mathbf{0} = (1, \ldots, 1). \tag{3.3}$$

We have seen in Section 2 that aging occurs only if the τ_x are sufficiently heavy-tailed. However, this certainly fails to be true for τ_x here: an easy calculation gives $\mathbb{P}[\tau_x \geq u] \leq u^{-\log u/2\beta^2 n}$ which decreases faster than any polynomial. It is therefore clear that if the process is given enough time to explore a large part of the configuration space and thus to discover the absence of heavy tails, then no aging occurs, at least in our picture. On the other hand, on shorter time scales the process does not feel the non-existence of heavy tails as can be seen from the following simple estimate. Let $\alpha > 0$, then

$$e^{\alpha^2 \beta^2 n/2} \mathbb{P}\big[\tau_x \geq u e^{\alpha \beta^2 n} (\alpha \beta \sqrt{2\pi n})^{-1/\alpha}\big]$$

$$= e^{\alpha^2 \beta^2 n/2} \mathbb{P}\Big[\frac{E_x}{\sqrt{n}} \geq \frac{\log u + \alpha \beta^2 n - \alpha^{-1} \log(\alpha \beta \sqrt{2\pi n})}{\beta \sqrt{n}}\Big] \tag{3.4}$$

$$\xrightarrow{n \to \infty} u^{-\alpha},$$

which can be obtained easily using $\mathbb{P}[E_x/\sqrt{n} \geq u] = (u\sqrt{2\pi})^{-1} e^{-u^2/2}(1+o(1))$. Therefore as $n \to \infty$

$$\mathbb{P}\Big[\frac{\tau_x}{e^{\alpha \beta^2 n}(\alpha \beta \sqrt{2\pi n})^{-1/\alpha}} \geq u\Big] = e^{-\alpha^2 \beta^2 n/2} \cdot u^{-\alpha}(1 + o(1)). \tag{3.5}$$

In view of the fact that the simple random walk on the hypercube almost never backtrack it seems reasonable to presume that if the process had time to make only approximately $e^{\alpha^2 \beta^2 n/2}$ steps, then aging could be observed. As we will see later this presumption shows to be true.

Before we state the aging result let us remark that there is a much stronger relation between "random exponentials" τ_x and heavy-tailed random variables. Let $(E_i, i \in \mathbb{N})$ be an i.i.d. sequence with the same common distribution as E_x. It was proved in [BBM05] that for some properly chosen $Z(n)$ and $N(n)$ the normalised sum

$$\frac{1}{Z(n)} \sum_{i=1}^{N(n)} e^{\beta \sqrt{n} E_i} \tag{3.6}$$

converges as $n \to \infty$ in law to an α-stable distribution with α depending on β and $N(n)$. Our methods allow to show that the same is true for the properly normalised clock process S_n, which is a properly-normalised sum of correlated random variables (see (1.7)).

We now fix objects for which we verify Conditions 1–6, or more precisely (A)–(D) together with Conditions 5, 6. The scales we choose should appear natural in view

of (3.5). We define

$$t(n) = g(n) = (\alpha\beta\sqrt{2\pi n})^{-1/\alpha} \exp(\alpha\beta^2 n), \tag{3.7}$$

$$r(n) = \rho(n)^{-1} = \exp(\alpha^2\beta^2 n/2), \tag{3.8}$$

$$\xi_n = mr(n), \tag{3.9}$$

$$T_\varepsilon^M(n, \alpha) = \left\{ x \in \mathcal{V}_n : \tau_x \in \big(\varepsilon g(n), Mg(n)\big) \right\}. \tag{3.10}$$

Theorem 3.1. *Let the parameters $\alpha \in (0, 1)$ and $\beta > 0$ be such that*

$$3/4 < \alpha^2\beta^2/2\log 2 < 1. \tag{3.11}$$

Then for $\mathbb{P}$-a.e. random environment $\boldsymbol{\tau}$

$$\lim_{n\to\infty} R_n(t(n), (1+\theta)t(n); \boldsymbol{\tau}) = \mathsf{Asl}_\alpha(1/1+\theta). \tag{3.12}$$

Remark. 1. We believe that the range of the validity (3.11) of the Theorem 3.1 is not the broadest possible. The upper bound 1 is correct. If $\alpha^2\beta^2/2\log 2 > 1$, then $\xi_n \gg 2^n$. That means that the state space $\mathcal{V}_n$ becomes too small and the process can feel its finiteness. On the other hand, the lower-bound $3/4$ is purely technical and can probably be improved. It appears because the potential-theoretic methods for the simple random walk on the hypercube that we use imply our conditions only if the set T_ε^M is sufficiently sparse, or more precisely, if it satisfies the so called minimal distance condition (see Proposition 3.2 and Lemma 3.7 below).

2. Observe that α is a free parameter. It is not fixed by the temperature β. Moreover, the condition (3.11) can be rewritten as

$$\alpha^{-1}\beta_c\sqrt{3/4} < \beta < \alpha^{-1}\beta_c, \tag{3.13}$$

where $\beta_c = \sqrt{2\log 2}$ is the critical temperature in the usual REM. This, in particular, means that aging can be observed in REM also above the critical temperature, $\beta < \beta_c$.

We now compare our results with those of [BBG03b]. To allow this comparison we describe very briefly some of the results of this paper. In [BBG03b] a discrete-time dynamics is considered, however as $n \to \infty$ this dynamics differs very little from the continuous-time dynamics used here. The most important object used to prove aging in [BBG03a] is a set of deep traps defined by $T_n(E) = \{x \in \mathcal{V}_n : E_x \geq u_n(E)\}$, where the function $u_n(E) = \beta_c\sqrt{n} + E/\beta_c\sqrt{n} + O(\sqrt{\log n/n})$ is chosen in the way that the distribution of $|T_n(E)|$ converges to some non-degenerate distribution as $n \to \infty$ and E is kept fixed. The mean of this distribution diverges if $E \to -\infty$ afterwards. The two-time function considered there is essentially the function $R_{T_n(E)}$ averaged over all the starting points in the top, let call it $\Pi_n(t_w, t_w + t, E)$. The main aging result of [BBG03b] says that for any $\beta > \beta_c$ and $\varepsilon > 0$

$$\lim_{t\to\infty}\lim_{E\to-\infty}\lim_{n\to\infty} \mathbb{P}\left[\left|\frac{\Pi_n(c_n t, (1+\theta)c_n t, E)}{\mathsf{Asl}_{\beta_c/\beta}(1/1+\theta)} - 1\right| > \varepsilon\right] = 0, \tag{3.14}$$

where $c_n \sim e^{\beta\sqrt{n}u_n(E)}$.

Here are the main differences between both results

1. Different two-time functions are considered. We believe that it is possible to eliminate the dependence on $T_n(E)$ (i.e. to convert something $R_{T_n(E)}$ to something R_n)

from (3.14) by a post-processing in the direction of Condition 5. It would be probably more difficult to get the a.s. convergence instead of the convergence in probability.

2. The main differences are in the considered top sizes and time scales. In [BBG03b] the size of the top is kept bounded as $n \to \infty$. This allows to apply "lumping techniques" to describe the properties of the projection of a simple random walk on the hypercube to the top, that is to prove that that an equivalent of our process U_n (see (2.6)) converges to the simple random walk on the complete graph with the vertex set $T_n(E)$. In our case the size of the top $T_\varepsilon^M(n)$ increases exponentially with n. This makes the application of the lumping more difficult. That is why we needed to develop different techniques to control the process U_n and in particular the random variables $\tau_{U_n(j)}$. These techniques can be found in Section 3.1. Using them we can verify the crucial Conditions (B) and (C) which imply Condition 3.

3. The time scale $c_n \sim e^{\beta\sqrt{n}u_n(E)} \sim e^{\beta\beta_c n + \beta E/\beta_c}$ used in [BBG03b] corresponds to the case $\alpha\beta/\beta_c = 1$ and is much larger than the scale $t(n) \sim e^{\alpha\beta^2 n} = e^{\frac{\alpha\beta}{\beta_c}\beta\beta_c n}$. These scales become closer if $\alpha\beta/\beta_c$ approaches 1, which is the upper limit of the validity of our theorem. We could probably, with some minor notational complications, improve our result to $t(n) = e^{\beta\beta_c n}h(n)$ with some $h(n) \to 0$ as $n \to \infty$ sufficiently fast, but even in this case $t(n) \ll c_n$. Another possibility would be to take the double limit as (3.14). This approach may be possibly practicable, however, it does not fall into our framework and we therefore prefer not to pursue it.

3.1. **Potential theory on the hypercube.**

In this section we study properties of the simple random walk on the hypercube, in particular hitting times of some relatively diluted but still large subsets of the hypercube.

Let us introduce some notation first. We write $I(x)$ for the rate function of the symmetric Bernoulli distribution on $\{0, 1\}$

$$I(x) := x \log x + (1 - x) \log(1 - x) + \log 2. \tag{3.15}$$

For $\gamma \in (1/2, 1)$ we use $\omega = \omega(\gamma)$ to denote the unique solution of the equation

$$I(\omega) = (2\gamma - 1) \log 2, \qquad 0 \leq \omega \leq 1/2. \tag{3.16}$$

Proposition 3.2. *Let for all $n \geq 1$ sets $A_n \subset \mathcal{V}_n$ be such that $|A_n| = \rho_n 2^n$, with "densities" ρ_n satisfying $\lim_{n\to\infty} \rho_n 2^{\gamma n} = \rho \in (0, \infty)$, for some $\gamma \in (1/2, 1)$. Let further the sets A_n satisfy the* minimal distance condition

$$\min\{d(x, y) : x, y \in A_n\} \geq (\omega(\gamma) + \varepsilon)n, \tag{3.17}$$

for some small constant $\varepsilon > 0$. Then for all $s \geq 0$

$$\lim_{n\to\infty} \max_{x\in A_n} \left| \mathbb{E}_x\left[\exp\left(-\frac{s}{2^{\gamma n}} H\left(A_n \setminus \{x\}\right)\right)\right] - \frac{\rho}{s + \rho}\right| = 0. \tag{3.18}$$

So that, the hitting time $H(A_n \setminus \{x\})/2^{\gamma n}$ is asymptotically exponentially distributed with mean $1/\rho$.

Remark. 1. Eventually, we will take A_n to be the set T_ε^M, that is a Poisson cloud on $\mathcal{V}_n$ with density $\rho(n)$. We will verify that the assumptions of the proposition are a.s. verified if γ is large enough. This is the result of Lemma 3.7 that can be found at the end of this section. 2. Note that in [BG06] similar results were obtained for sets with the minimal

distance between the points $o(n)$, but, on the other hand, the size of these sets should be much smaller than in our case, namely $O(\log n)$.

Proof. We will use a method introduced by Matthews [Mat88] to show this Proposition. The advantage of this method is that the hitting time of a large set can be controlled by means of much simple estimates on the hitting time of a point.

We define

$$f_n^+(s) := \max\{\mathbb{E}_x[\exp(-s2^{-\gamma n}H(y))] : x, y \in A_n, x \neq y\},$$
$$f_n^-(s) := \min\{\mathbb{E}_x[\exp(-s2^{-\gamma n}H(y))] : x, y \in A_n, x \neq y\} \tag{3.19}$$

The following lemma is the most important building block of the proof.

Lemma 3.3 (Theorem 1.3 of [Mat88]). *For any $x \in A_n$*

$$\frac{\Gamma(1/f_n^-(s))}{\Gamma(1/f_n^+(s))} \cdot \frac{\Gamma(|A_n|)}{\Gamma(|A_n|-1)} \cdot \frac{\Gamma(|A_n|-2+1/f_n^+(s))}{\Gamma(|A_n|-1+1/f_n^-(s))}$$

$$\leq \mathbb{E}_x\left[e^{-sH(A_n\setminus\{x\})/2^{\gamma n}}\right] \tag{3.20}$$

$$\leq \frac{\Gamma(1/f_n^+(s))}{\Gamma(1/f_n^-(s))} \cdot \frac{\Gamma(|A_n|)}{\Gamma(|A_n|-1)} \cdot \frac{\Gamma(|A_n|-2+1/f_n^-(s))}{\Gamma(|A_n|-1+1/f_n^+(s))},$$

where Γ is the usual gamma-function.

Proof. The lemma follows directly from Theorem 1.3 of [Mat88]. Using hat for objects as they appear in that paper, we identify $\hat{N}$ with $|A_n| - 1$ and $\hat{A}_i$, $i = 1, \ldots, \hat{N}$ with points of $A_n \setminus \{x\}$. Since we are interested in the first visit of the set $A_n \setminus \{x\}$, the expression (3.20) is obtained by setting $\hat{n} = 1$ in Matthews' theorem, and by rewriting the products appearing there using Γ functions. $\qquad\square$

To apply the previous lemma we need very precise estimates on f_n^+ and f_n^-. Later in this section we will show

Lemma 3.4. *Let the assumptions of Proposition 3.2 be satisfied. Then for all $s > 0$ the functions f_n^+ and f_n^- satisfy*

$$\lim_{n\to\infty} \frac{1}{2^{(1-\gamma)n}f_n^+(s)} = \lim_{n\to\infty} \frac{1}{2^{(1-\gamma)n}f_n^-(s)} = s, \tag{3.21}$$

$$\lim_{n\to\infty} \frac{1}{f_n^-(s)} - \frac{1}{f_n^+(s)} = 0. \tag{3.22}$$

To finish the proof of Proposition 3.2 we will need another elementary technical lemma.

Lemma 3.5. *Let a_m, b_m, δ_m be such that $a_m \to a$, $b_m \to b$ and $\delta_m \to 0$ as $m \to \infty$ with $0 < a, b < \infty$. Then*

$$\lim_{m\to\infty} \frac{\Gamma(a_m m)}{\Gamma(a_m m + \delta_m)} \cdot \frac{\Gamma(b_m m + \delta_m)}{\Gamma(b_m m)} = 1. \tag{3.23}$$

Proof. By Stirling formula, $\Gamma(m) = \sqrt{2\pi}(m-1)^{m-\frac{1}{2}}e^{-m+1}(1+o(1))$. Therefore, up to a multiplicative correction $1 + o(1)$, the expression inside the limit (3.23) equals

$$\frac{\left(1 + \frac{\delta_m}{b_m m - 1}\right)^{b_m m - \frac{1}{2}}}{\left(1 + \frac{\delta_m}{a_m m - 1}\right)^{a_m m - \frac{1}{2}}} \cdot \frac{(b_m m - 1 + \delta_m)^{\delta_m}}{(a_m m - 1 + \delta_m)^{\delta_m}}. \tag{3.24}$$

The logarithm of the numerator of the first fraction satisfies

$$\left(b_m m - \frac{1}{2}\right) \log\left(1 + \frac{\delta_m}{b_m m - 1}\right) \le c\delta_m \xrightarrow{m\to\infty} 0. \tag{3.25}$$

The same hold for the denominator. The logarithm of the second fraction is

$$\delta_m \log \frac{b_m m - 1 + \delta_m}{a_m m - 1 + \delta_m} \xrightarrow{m\to\infty} 0. \tag{3.26}$$

This finishes the proof. $\qquad\square$

We now use Lemmas 3.3–3.5 to finish the proof of Proposition 3.2. Define

$$\begin{aligned}
m = m(n) &= 2^{(1-\gamma)n} & a_m &= 2^{-(1-\gamma)n}(|A_n| - 2 + 1/f_n^+(s)) \\
\delta_m &= 1/f_n^-(s) - 1/f_n^+(s) & b_m &= 2^{-(1-\gamma)n}(1/f_n^+(s))
\end{aligned} \tag{3.27}$$

Then, by Lemma 3.4, a_m, b_m, δ_m satisfy the assumptions of Lemma 3.5. Therefore,

$$\lim_{n\to\infty} \frac{\Gamma(1/f_n^-(s))}{\Gamma(1/f_n^+(s))} \cdot \frac{\Gamma(|A_n| - 2 + 1/f_n^+(s))}{\Gamma(|A_n| - 2 + 1/f_n^-(s))} = 1. \tag{3.28}$$

Comparing the last display with the left-hand side of (3.20) we get using Lemma 3.4

$$\mathbb{E}_x\left[e^{-s2^{-\gamma n}H(A_n\setminus\{x\})}\right] \ge \frac{|A_n| - 1}{|A_n| - 2 + 1/f_n^-(s)}(1 + o(1)) = \frac{\rho}{s + \rho}(1 + o(1)). \tag{3.29}$$

The corresponding upper bound can be obtained analogously. $\qquad\square$

Proof of Lemma 3.4. To estimate $f_n^+(s)$ and $f_n^-(s)$ we need to compute the Laplace transform $f_n(x, y; s) := \mathbb{E}_x[\exp(-s2^{-\gamma n}H(y))]$ for all pairs $x, y \in A_n$. This task can be largely simplified using symmetries of the hypercube. First, it is obvious that $f_n(x, y; s)$ depends only on the distance between x and y, that is if $d(x, y) = k$, then $f_n(x, y; s) = f_n(k, s)$, where

$$f_n(k, s) := \mathbb{E}_{z_k}[\exp(-s2^{-\gamma n}H(\mathbf{0}))], \qquad z_k := (\overbrace{-1, \ldots, -1}^{k \text{ times}}, 1, \ldots, 1). \tag{3.30}$$

Second, any walk started at z_k should visit a point in the distance $l < k$ from $\mathbf{0}$ before hitting $\mathbf{0}$. This implies that $f_n(k, s) \le f_n(l, s)$. Therefore, using the minimal distance condition (3.17),

$$f_n(n, s) \le f_n^-(s) \le f_n^+(s) \le f_n((\omega + \varepsilon)n, s). \tag{3.31}$$

The statement of Lemma 3.4 is thus equivalent to

$$\lim_{n\to\infty} \frac{1}{2^{(1-\gamma)n}f_n(n, s)} = s, \tag{3.32}$$

$$f_n((\omega + \varepsilon)n, s) - f_n(n, s) = o(2^{-2(1-\gamma)n}). \tag{3.33}$$

We first compute $f_n(n, s) = \mathbb{E}_{z_n}\left[e^{-s2^{-\gamma n}H(\mathbf{0})}\right]$. This computation closely follows Section 3 of [Mat89]. We should be a little bit more careful, since we need to compute the Laplace transform of $H(\mathbf{0})$ on a scale that is not typical for it, normally $H(\mathbf{0}) \sim 2^n$. (In

[Mat89] $\mathbb{E}_x\big[e^{-s2^{-n}H(\mathbf{0})}\big]$ was estimated.) By Fourier methods for random walks on finite groups ([Mat88, Mat89], see also [Dia88] for the general theory)

$$\mathbb{E}_x e^{-\lambda H(\mathbf{0})} = \frac{\sum_{y \in \mathcal{V}_n}(-1)^{x \cdot y}\big[1 - e^{-\lambda}(1 - \frac{2d(y,\mathbf{0})}{n})\big]^{-1}}{\sum_{y \in \mathcal{V}_n}\big[1 - e^{-\lambda}(1 - \frac{2d(y,\mathbf{0})}{n})\big]^{-1}}, \tag{3.34}$$

where $x \cdot y = \sum_{i=1}^{n} x_i y_i$ is the standard scalar product in $\mathbb{R}^n$. Setting $\lambda = s/2^{\gamma n}$, the denominator of (3.34) (which does not depend on x) equals

$$\sum_{i=0}^{n} \binom{n}{i}\Big[1 - e^{-s/2^{\gamma n}}\Big(1 - \frac{2i}{n}\Big)\Big]^{-1}. \tag{3.35}$$

We expand $e^{-s/2^{\gamma n}} = 1 - s/2^{\gamma n} + O(2^{-2\gamma n})$. Treating separately the term $i = 0$, the denominator becomes

$$(1 - e^{-s/2^{\gamma n}})^{-1}\Big\{1 + \sum_{i=1}^{n}\frac{s}{2^{\gamma n}} \cdot \frac{n}{2i}\binom{n}{i}\Big\}(1 + o(1)). \tag{3.36}$$

Writing $\frac{n}{2i} = \sum_{j=0}^{\infty}(1 - \frac{2i}{n})^j$ for $i \notin \{0, n\}$, we get

$$\sum_{i=1}^{n}\frac{s}{2^{\gamma n}} \cdot \frac{n}{2i}\binom{n}{i} = \frac{s}{2^{\gamma n}}\sum_{j=0}^{\infty}\Big(\frac{2}{n}\Big)^j \sum_{i=1}^{n-1}\binom{n}{i}\Big(\frac{n}{2} - i\Big)^j + O(2^{-\gamma n}) \tag{3.37}$$

$$= s2^{(1-\gamma)n}(1 + O(n^{-1})).$$

To evaluate the inner sum we used the fact that it is, after the normalisation by 2^{-n}, up to a small error, equal to the j-th central moment of the binomial distribution with parameters n and $1/2$. The denominator of (3.34) is therefore equal to

$$s2^{(1-\gamma)n}(1 - e^{-s/2^{\gamma n}})^{-1}(1 + o(1)) = 2^n(1 + o(1)). \tag{3.38}$$

The calculation of the numerator of (3.34) for $x = z_n$ can be done analogously. The only difference is that all factors $\binom{n}{i}$ in (3.35) and further should be replaced by $C_n(i) = (-1)^i\binom{n}{i}$. We need therefore to compute

$$\sum_{i=1}^{n} C_n(i)\frac{s}{2^{\gamma n}} \cdot \frac{n}{2i}, \tag{3.39}$$

which corresponds to the sum in (3.36).

Lemma 3.6.

$$\sum_{i=1}^{n}(-1)^i\binom{n}{i}\frac{1}{i} = -1 - \frac{1}{2} - \cdots - \frac{1}{n}. \tag{3.40}$$

Proof. It is easy to see that

$$\frac{\mathrm{d}}{\mathrm{d}x}\big(x^{-1}[(1 - x)^n - 1]\big) = \sum_{i=1}^{n}(-1)^i\binom{n}{i}\frac{x^i}{i}. \tag{3.41}$$

Therefore,

$$\sum_{i=1}^{n}(-1)^i\binom{n}{i}\frac{1}{i} = \int_0^1\frac{(1 - x)^n - 1}{x}\,\mathrm{d}x = \int_0^1\frac{v^n - 1}{1 - v}\,\mathrm{d}v \tag{3.42}$$

from which it is easy to finish the proof. $\square$

From the last lemma it follows easily that (3.39) tends to 0 as $n \to \infty$, and the numerator of (3.34) for $x = z_n$ is equal to

$$(1 - e^{-s/2^{\gamma n}})^{-1}(1 + o(1)) = \frac{2^{\gamma n}}{s}(1 + o(1)). \tag{3.43}$$

Putting together (3.34), (3.38), and (3.43) we get

$$f_n(n, s) = \mathbb{E}_{z_n}\left[e^{-s2^{-\gamma n}H(0)}\right] = \frac{1}{s2^{(1-\gamma)n}}(1 + o(1)) \qquad \text{as } n \to \infty. \tag{3.44}$$

This proves (3.32).

It remains to prove (3.33). We use again formula (3.34). Since the denominator does not depend on x, we should only compute the numerator of the difference $f_n(h, s) - f_n(n, s)$ (eventually we will take $h = (\omega + \varepsilon)n$ as in (3.17)). Observing that the terms with $d(\mathbf{0}, y) = 0$ cancel, the numerator of the difference is equal to

$$\sum_{i=1}^{n} D_n(i)\left[1 - e^{-s/2^{\gamma n}}\left(1 - \frac{2i}{n}\right)\right]^{-1}, \tag{3.45}$$

where the combinatorial factors $D_n(i)$ are given by

$$D_n(i) = \sum_{y:d(\mathbf{0},y)=i} \left[(-1)^{z_h \cdot y} - (-1)^i\right]. \tag{3.46}$$

An easy combinatorial reasoning gives that (using $\binom{n}{k} := 0$ for $k \notin \{0, \ldots, n\}$)

$$D_n(i) = -2(-1)^i \sum_{j=1,3,5,\ldots}^{n-h} \binom{n-h}{j}\binom{h}{i-j}. \tag{3.47}$$

Hence, expanding again $e^{-s/2^{\gamma n}}$, (3.45) can be written as

$$-\sum_{i=1}^{n} 2(-1)^i \sum_{j=1,3,5,\ldots}^{n-h} \binom{n-h}{j}\binom{h}{i-j}\frac{n}{2i}(1 + o(1))$$

$$= -n \sum_{j=1,3,5,\ldots}^{n-h} \sum_{k=0}^{h} (-1)^{j+k}\binom{n-h}{j}\binom{h}{k}\frac{1}{j+k}(1 + o(1)). \tag{3.48}$$

The sum over k can be computed similarly as in Lemma 3.6: for $j \in \mathbb{N}$

$$\sum_{k=0}^{h} \frac{(-1)^k}{j+k}\binom{h}{k} = \int_0^1 x^{j-1}(1 - x)^h \, dx = \frac{\Gamma(1+h)\Gamma(j)}{\Gamma(1+h+j)} = \frac{1}{j}\binom{h+j}{j}^{-1}. \tag{3.49}$$

Observing that $(-1)^j = -1$, the expression (3.45) equals (up to a multiplicative correction $1 + o(1)$)

$$n \sum_{j=1,3,5,\ldots}^{n-h} \binom{n-h}{j}\binom{h+j}{h}^{-1}\frac{1}{j} =: n \sum_{j=1,3,5,\ldots}^{n-h} K(j). \tag{3.50}$$

Taking $h = \lfloor \xi n \rfloor$ for some $\xi \in (0, 1/2)$, and calculating the ratio of the consecutive summands $K(j)/K(j+2)$, we find that $K(j)$ attains a maximum for $j \sim n(\frac{1}{2} - \xi)$. The last display is therefore bounded from above by

$$Cn^2 \binom{(1-\xi)n}{(\frac{1}{2}-\xi)n} \binom{\frac{1}{2}n}{\xi n}^{-1}. \tag{3.51}$$

Standard Stirling type estimates applied to the previous expression imply that the numerator of the difference $f_n(\xi n, s) - f_n(n, s)$ is, for all n large enough, bounded by

$$\exp[n(I(\xi) + \delta)] = 2^{nI(\xi - \delta')/\log 2}, \tag{3.52}$$

where $\delta, \delta' > 0$ can be taken arbitrarily small. Taking now $\xi = \omega + \varepsilon$, choosing $\delta' < \varepsilon$, using (3.17), the fact that the denominator is of order 2^n (see (3.38)), and $I(x)$ is decreasing on $[0, 1/2]$, it is easy to prove (3.33). This completes the proof of Lemma 3.4. $\square$

The next lemma specifies conditions under which the assumptions of Proposition 3.2 are verified for a Poisson cloud on the hypercube. It also collects some technical facts about this cloud that we will need later.

Lemma 3.7. *Let A_n be a sequence of Poisson clouds on $\mathcal{V}_n$ with densities ρ_n satisfying $\lim_{n \to \infty} \rho_n 2^{\gamma n} = \rho \in (0, \infty)$ for some $\gamma \in (0, 1)$. Then*
(i) $\mathbb{P}$-a.s. $2^{(\gamma - 1)n} |A_n| \in (\rho - \delta, \rho + \delta)$ for any $\delta > 0$ and n large enough.
(ii) Let $\gamma > 1/2$ and let $\omega' < 1/2$ be the unique solution of $I(\omega') = 2(1 - \gamma) \log 2$. Then $\mathbb{P}$-a.s.

$$d_{\min} := \min\{d(x, y) : x, y \in A_n\} \geq (\omega' - \delta)n \tag{3.53}$$

for any $\delta > 0$ and n large enough.
(iii) If $\gamma > 3/4$, then the set A_n satisfies $\mathbb{P}$-a.s. the assumptions of Proposition 3.2 for all n large enough.
(iv) The claims (ii), (iii) stay valid if the set A_n is replaced by $A_n \cup \{0\}$.

As a corollary of Proposition 3.2 and Lemma 3.7 we get

Corollary 3.8. *If $\alpha^2 \beta^2 / 2 \log 2 > 3/4$ then Condition (C) holds for $r(n)$ and $\rho(n)$ as defined in (3.8) with $\mathcal{K}_r = 1$.*

Proof of Lemma 3.7. (i) The proof is standard. Using the Chebyshev inequality we get

$$\mathbb{P}\big[|2^{(\gamma-1)n}|A_n| - \rho| \geq \delta\big] \leq \delta^{-2} 2^{n(\gamma - 1)}. \tag{3.54}$$

The Borel-Cantelli lemma then implies the result.
(ii) We construct the set A_n in the following way. Let R be a binomial random variable with parameters 2^n and ρ_n, and let $(x_i, i = 1, \ldots, R)$ be a collection of randomly chosen points in $\mathcal{V}_n$, such that given $(x_j, j < i)$ the point x_i is uniformly distributed in $\mathcal{V}_n \setminus \{x_j : j < i\}$. It is easy to estimate the probability that x_i is too close to some of x_j, $j < i$. Indeed, the volume of the ball with radius $(\omega' - \delta)n$ around a point x satisfies

$$\big|\{y \in \mathcal{V}_n : d(x, y) \leq (\omega' - \delta)n\}\big| \leq e^{n(I(0) - I(\omega' - \delta) + \varepsilon)} \tag{3.55}$$

for all $\varepsilon > 0$ and n large enough. Therefore, for all $i \leq R$

$$\mathbb{P}\big[\min_{j < i} d(x_i, x_j) \leq (\omega' - \delta)n\big] \leq \frac{i - 1}{2^n - i + 1} e^{n(I(0) - I(\omega' - \delta) + \varepsilon)}. \tag{3.56}$$

So that, summing over i,

$$\mathbb{P}\big[d_{\min} \le (\omega' - \delta)n\big] \le \sum_{r=1}^{2^n} \mathbb{P}[R = r] \sum_{i=1}^{r} \mathbb{P}\Big[\min_{j<i} d(x_i, x_j) \le (\omega' - \delta)n\Big]$$

$$\le \mathbb{P}\big[R \ge (1 + \varepsilon)2^{(1-\gamma)n}\big] + (1 + \varepsilon)^2 2^{2(1-\gamma)n} e^{n(-I(\omega'-\delta)+\varepsilon)}. \quad (3.57)$$

It follows from the assumptions of the lemma that for any δ we can find $\varepsilon > 0$ such that the right-hand side of the last expression decays exponentially with n. The Borel-Cantelli lemma then implies (ii).

(iii) The assumption of Proposition 3.2 are satisfied if $d_{\min} \ge (\omega + \varepsilon)n$ for some $\varepsilon > 0$ and ω defined in (3.16). Using part (ii) of this lemma, this condition is satisfied if $\omega' - \delta \ge \omega + \varepsilon$ for some ε, δ. Since $I(\cdot)$ is decreasing on $(0, 1/2)$, we get using the definitions of ω, ω' that γ should satisfy $2(1 - \gamma) < 2\gamma - 1$. This is true for any $\gamma > 3/4$.

Easy modifications of the proofs of (ii), (iii) to get (iv) are left to the reader. $\square$

3.2. **Proof of Theorem 3.1.** We can now finish the proof of Theorem 3.1. Since Condition (A) is trivially verified and (C) follows from Corollary 3.8, it remains to verify Conditions (B), (D), 5 and 6 for the choice (3.7)–(3.10) of the parameters.

We need one technical lemma first:

Lemma 3.9. *(a) Green's function $G^n_{\xi_n}$ satisfies (recall $\xi_n = mr(n)$)*

$$\limsup_{n\to\infty} G^n_{\xi_n}(\mathbf{0}, \mathbf{0}) = \limsup_{n\to\infty} G^n_{\xi_n \log \xi_n}(\mathbf{0}, \mathbf{0}) =: G_\infty = 1. \quad (3.58)$$

(b) For all $x \ne \mathbf{0}$ and for all n large enough

$$G^n_{\xi_n}(\mathbf{0}, x) \le cn^{-1}. \quad (3.59)$$

Proof. (a) Since $G^n_{\xi_n}(x, y) \le G^n_{\xi_n \log \xi_n}(x, y)$ it is sufficient to consider only the second limit. The probability that Y_n returns to $\mathbf{0}$ before $\xi_n \log \xi_n$ tends to 0 as $n \to \infty$. Indeed, using the Chebyshev inequality,

$$\mathbb{P}_{z_1}[H(\mathbf{0}) < \xi_n \log \xi_n] \le e^{s\xi_n \log \xi_n / 2^n} \mathbb{E}_{z_1}\big[e^{-sH(\mathbf{0})/2^n}\big]. \quad (3.60)$$

Performing more carefully the same computation as the one on page 138 of [Mat89] it can be proved that

$$\mathbb{E}_{z_1}\big[e^{-sH(\mathbf{0})/2^n}\big] = \frac{1 + s/n\big[1 + O(s2^{-n}) + O(1/n)\big]}{1 + s[1 + O(s2^{-n})][1 + 1/n + O(n^{-2})]}. \quad (3.61)$$

Take now $s = s(n)$ such that $s(n) \to \infty$ and $s(n)\xi_n \log \xi_n / 2^n \to 0$ as $n \to \infty$, which is possible by assumption (3.11) of Theorem 3.1. Then the right-hand side of (3.60) tends to 0 as $n \to \infty$. Therefore, for any $\varepsilon > 0$ there exists n_0 such that for all $n > n_0$ $\mathbb{P}_{z_1}[H(\mathbf{0}) \le \xi_n \log \xi_n] \le \varepsilon$. Since after every visit of $\mathbf{0}$ the random walk Y_n should jump to some of its neighbours, we get by iterating the last estimate

$$\mathbb{P}\Big[\sum_{i=0}^{\xi_n \log \xi_n} \mathbb{1}\{Y_n = \mathbf{0}\} \ge k\Big] \le \varepsilon^{k-1}. \quad (3.62)$$

Therefore

$$G^n_{\xi_n}(\mathbf{0}, \mathbf{0}) = \mathbb{E}\left[\sum_{i=0}^{\xi_n \log \xi_n} \mathbb{1}\{Y_n = \mathbf{0}\}\right] \le (1 - \varepsilon)^{-1} \tag{3.63}$$

for all large n. Since ε was arbitrary, the proof of (a) is finished.

(b) Since $G^n_{\xi_n}(\mathbf{0}, x) \le \mathbb{P}[H(x) \le \xi_n] G^n_{\xi_n}(\mathbf{0}, \mathbf{0})$, and $\mathbb{P}[H(x) \le \xi_n]$ decreases as $\mathrm{dist}(x, \mathbf{0})$ increases, it is sufficient to show that $\mathbb{P}[H(z_1) \le \xi_n] = \mathbb{P}_{z_1}[H(\mathbf{0}) \le \xi_n] \le cn^{-1}$. However, this is a direct consequence (3.61) and Chebyshev inequality with $s(n) = n$. □

Condition (D). We verify this condition with $\lambda_n = n^{1/2}$. Then,

$$\sum_{n=0}^{\infty} \exp\{-\lambda_n t(n)/g(n)\} = \sum_{n=0}^{\infty} \exp\{-\lambda_n\} < \infty \tag{3.64}$$

as required. Using the last lemma we get

$$\sum_{x \in \mathcal{V}_n} (e^{\lambda_n G^n_{\xi_n}(\mathbf{0}, x)} - 1) = e^{2n^{1/2}} + \sum_{x \in \mathcal{V}_n \setminus \{\mathbf{0}\}} (e^{\lambda_n G^n_{\xi_n}(\mathbf{0}, x)} - 1). \tag{3.65}$$

For $x \in \mathcal{V}_n \setminus \{\mathbf{0}\}$ Lemma 3.9 yields that $\lambda_n G^n_{\xi_n}(\mathbf{0}, x) \le cn^{-1/2}$. Since $e^u - 1 \le 2u$ for all u sufficiently close to 0, the last display is bounded from above by

$$e^{2n^{1/2}} + \sum_{x \in \mathcal{V}_n \setminus \{\mathbf{0}\}} 2G^n_{\xi_n}(\mathbf{0}, x) \le e^{2n^{1/2}} + 2\xi(n) \le Cmr(n), \tag{3.66}$$

which was to be proved.

Condition (B). This condition follows from the next lemma.

Lemma 3.10. *Uniformly for* $x \in T^M_\varepsilon$

$$\lim_{n \to \infty} G^n_{T^M_\varepsilon \setminus \{x\}}(x, x) = 1. \tag{3.67}$$

Proof. Let $H'_n(x) = \min\{i \ge 1 : Y_n(i) = x\}$. Then

$$G^n_{T^M_\varepsilon \setminus \{x\}}(x, x) = \left(\mathbb{P}_x[H_n(T^M_\varepsilon \setminus \{x\}) < H'_n(x)]\right)^{-1} \tag{3.68}$$

However

$$1 - \mathbb{P}_x[H_n(T^M_\varepsilon \setminus \{x\}) < H'_n(x)]$$
$$\le \mathbb{P}_x[H'_n(x) \le \xi_n \log \xi_n] + \mathbb{P}_x[H_n(T^M_\varepsilon \setminus \{x\}) \ge \xi_n \log \xi_n]. \tag{3.69}$$

The first term is independent of x and converges to 0, as can be proved e.g. using Lemma 3.9(a). Since $H_n(T^M_\varepsilon \setminus \{x\})/2^{\gamma n}$ is asymptotically exponentially distributed with mean $1/p^M_\varepsilon$ as follows from Corollary 3.8, and since $\xi_n = m2^{\gamma n}$, the second term in (3.69) converges also to 0 uniformly in x. The lemma then follows from (3.68) and (3.69). □

Condition 5. Let t_n be a deterministic time sequence and A_n defined as in (2.13). Fix ε small enough such that $h(\varepsilon)$ used in Condition 1 satisfies $h(\varepsilon) \le \delta^2/4$. Then by Chebyshev inequality the time spent in the shallow traps is small,

$$\mathbb{P}\left[\sum_{i=0}^{\xi_n} e_i \tau_{Y_n(i)} \mathbb{1}\{Y_n(i) \in T^\varepsilon(n)\} \ge t(n)\delta/2 \,\Big|\, \tau\right] \le \delta/2. \tag{3.70}$$

Let B be the event from the previous display. Conditionally on B and A_n, the event $\{X_n(t_n) \neq U_n(j_n)\}$ occurs only if X_n returns to $U_n(j_n)$ before the time $S_n(r_n(j+1))$. However, the probability of such return tends to 0 by Lemma 3.9(a). This finishes the proof.

Condition 6. As follows from (C) the number ζ_n of visited deep traps converges to Poisson distribution with mean $\mathcal{K}_r m p_\varepsilon^M$. The probability that Y_n returns to any of these traps before ξ_n converges to 0 by Lemma 3.9(a). From these two facts the condition follows easily.

4. Aging on the torus

As a second application of the general strategy presented in Section 2 we will give a proof of aging on a two-dimensional torus. This complements the results of [BČM06] about aging on $\mathbb{Z}^2$. Note that aging on d-dimensional torus, $d \geq 3$, could be proved analogously, however for notational convenience we treat only the two-dimensional case.

We will consider the following model. Let $G_n = (\mathcal{V}_n, \mathcal{E}_n)$ be a two-dimensional torus of size 2^n with nearest-neighbours connection, i.e. $\mathcal{V}_n = \mathbb{Z}^2/2^n\mathbb{Z}^2$, and edge $\langle x, y \rangle \in \mathcal{E}_n$ iff $\sum_{i=1}^2 |x_i - y_i| \bmod 2^n = 1$. We use $d(x, y)$ to denote the graph distance of $x, y \in \mathcal{V}_n$. Let further $\boldsymbol{\tau} = \{\tau_x^n\}$, $x \in \mathcal{V}_n$, $n \in \mathbb{N}$, be a collection of positive i.i.d. random variables satisfying

$$\mathbb{P}[\tau_x^n \geq u] = u^{-\alpha}(1 + o(1)) \qquad (u \to \infty). \tag{4.1}$$

For simplicity we assume that $\mathbb{P}[\tau_x \geq 1] = 1$. Given graph G_n and the random environment $\boldsymbol{\tau}$ we will consider the Markov processes X_n defined in Section 1. For this process we will show:

Theorem 4.1. *Let $t(n) = 2^{2n/\alpha}n^{1-(\gamma/\alpha)}$ for some $\gamma \in (0, 1/6)$. Then for $\mathbb{P}$-a.e. realisation of the random environment $\boldsymbol{\tau}$*

$$\lim_{n\to\infty} R_n(t(n), (1+\theta)t(n); \boldsymbol{\tau}) = \mathsf{Asl}_\alpha(1/1+\theta). \tag{4.2}$$

To show Theorem 4.1 we will verify that Conditions (A)–(D), 5, and 6 hold in our case for the parameters

$$t(n) = 2^{2n/\alpha}n^{1-\gamma/\alpha}, \qquad \xi_n = mr(n) = m2^{2n}n^{1-\gamma},$$
$$g(n) = \rho(n)^{-1/\alpha} = 2^{2n/\alpha}n^{-\gamma/\alpha}, \qquad f(n) = n, \tag{4.3}$$
$$T_\varepsilon^M(n) = \{x \in \mathcal{V}_n : \tau_x \in (\varepsilon g(n), Mg(n))\}.$$

The main motivation for Theorem 4.1 is to extend the range of aging scales on $\mathbb{Z}^2$ and mainly to really explore the extreme values of the random landscape. Namely, the BTM on the whole lattice $\mathbb{Z}^2$ does not find the deepest traps that are close to its starting position. In the first 2^{2n} steps, it gets to the distance 2^n and visits $O(2^{2n}/\log(2^{2n})) = O(2^{2n}/n)$ sites. Therefore, the deepest visited trap has a depth of order $2^{2n/\alpha}/n^{1/\alpha}$, which is much smaller that the depth of the deepest trap in the disk with radius 2^n, that is $2^{2n/\alpha}$. Eventually, the process visits also this deepest trap, however it will be too late. This trap will no longer be relevant for the time change since much deeper traps will have already be visited. The deepest trap is relevant only if the random walk stays in the neighbourhood of its starting point a long enough time. One way to force it to stay

is to change $\mathbb{Z}^2$ to the torus. By changing the size of the torus relatively to the number of considered steps, i.e. by changing γ, different depth scales become relevant for aging.

The range of possible values of $\gamma \in (0, 1/6)$ has, as in the REM case, a natural bound and an artificial one. It is natural that $\gamma < 0$ cannot be considered, since in this case the number of steps ξ_n of the simple walk is larger than $2^{2n}(\log 2^n)^{1+\varepsilon}$, $\varepsilon > 0$. Therefore, its occupation probabilities are very close to the uniform measure on the torus, that is the process is almost in equilibrium. The other bound, $\gamma = 1/6$, comes from the techniques that we use. We do not believe it to be meaningful since we expect the theorem to hold for all $\gamma > 0$. Actually, the result for $\gamma > 1$ follows easily from the result of [BČM06] for the whole lattice. In this case the size of the torus is much larger than ξ_n^2. So that, the process has no time to discover the difference between the torus and $\mathbb{Z}^2$. We also know that Theorem 4.1 holds also in the window $[1/6, 1]$ since it can be proved by the same methods as for $\mathbb{Z}^2$, [BČM06]. However, to keep the presentation in this paper compact and to avoid the more complicated coarse-graining techniques of [BČM06], we prefer to stick to $\gamma \in (0, 1/6)$.

To show our conditions we need again a little bit of the potential theory for the simple random walk Y_n on the torus G_n.

4.1. Potential theory on the torus.
The following proposition is an equivalent of Proposition 3.2.

Proposition 4.2. *Let $A_n \subset \mathcal{V}_n$ be such that $|A_n| = \rho_n 2^{2n}$ with the "density" ρ_n satisfying $\lim_{n \to \infty} 2^{2n} n^{-\gamma} \rho_n = \rho$ for some $\gamma \in (0, 1/6)$ and $\rho \in (0, \infty)$. Let further A_n satisfy the minimal distance condition*

$$\min\{d(x, y) : x, y \in A_n\} \geq 2^n n^{-\kappa} =: r_{\min}, \tag{4.4}$$

for some $\kappa > 0$. Then, for $\mathcal{K}_r = \pi(2 \log 2)^{-1}$,

$$\lim_{n \to \infty} \max_{x \in A_n} \left| \mathbb{E}_x \left[\exp\left(-\frac{s}{2^{2n} n^{1-\gamma}} H(A_n \setminus \{x\}) \right) \right] - \frac{\mathcal{K}_r \rho}{s + \mathcal{K}_r \rho} \right| = 0. \tag{4.5}$$

Proof. To proof this proposition we use again the methods from [Mat88], namely the formula (3.20). To apply it, we need to get precise estimates of the following functions (we use translation invariance of the torus and the minimal distance condition to simplify the definitions)

$$\begin{aligned}
f_n^+(s) &= \sup_{x \in \mathcal{V}_n^{\text{out}}} \mathbb{E}_x \left[\exp\left(-\frac{s}{2^{2n} n^{1-\gamma}} H(\mathbf{0}) \right) \right] \\
f_n^-(s) &= \inf_{x \in \mathcal{V}_n^{\text{out}}} \mathbb{E}_x \left[\exp\left(-\frac{s}{2^{2n} n^{1-\gamma}} H(\mathbf{0}) \right) \right],
\end{aligned} \tag{4.6}$$

where $\mathcal{V}_n^{\text{out}} = \{x \in \mathcal{V}_n : d(\mathbf{0}, x) \geq r_{\min}\}$, and $\mathbf{0} = (0, 0)$. Using the same reasoning as between (3.20) and (3.29) it is easy to show that Proposition 4.2 follows from

$$\lim_{n \to \infty} \frac{1}{n^\gamma f_n^{\pm}(s)} = \mathcal{K}_r^{-1} s, \tag{4.7}$$

$$\lim_{n \to \infty} \left(\frac{1}{f_n^-(s)} - \frac{1}{f_n^+(s)} \right) = 0. \tag{4.8}$$

To estimate the functions f_n^+ and f_n^- we use methods developed in [Cox89, CD02]. Following these papers we denote by $q_k(x, y)$ for the (smoothed) probability that a simple random walk on $\mathbb{Z}^2$ started at x is at y after k steps, more precisely we define

$$q_k(x, y) := \{\mathbb{P}_x[Y(k) = y] + \mathbb{P}_x[Y(k + 1) = y]\}/2 \qquad x, y \in \mathbb{Z}^2, \tag{4.9}$$

where Y is a simple random walk on $\mathbb{Z}^2$. Let $L := 2^n$. We use

$$q_k^n(x, y) := \sum_{z \in \mathbb{Z}^2} q_k(x, y + zL), \qquad x, y \in \mathcal{V}_n, \tag{4.10}$$

to denote the transition probability of the simple random walk on the torus $\mathcal{V}_n$. Let

$$G^n(x; s) = \sum_{k=0}^{\infty} e^{-sk/h(n)} q_k^n(\mathbf{0}, x) \tag{4.11}$$

with $h = h(n) = 2^{2n} n^{1-\gamma}$. As is shown e.g. in [Cox89], p. 1339,

$$f_n^{\pm}(s) = \sup_{x \in \mathcal{V}_n^{\mathrm{out}}} \left(\inf_{x \in \mathcal{V}_n^{\mathrm{out}}} \right) \frac{G^n(x; s)}{G^n(\mathbf{0}; s)}. \tag{4.12}$$

Adapting slightly the calculation in [Cox89], p. 1340, we find that the common denominator in the previous formula satisfies as $n \to \infty$

$$G^n(\mathbf{0}; s) = \frac{2n}{\pi} \log 2(1 + o(1)). \tag{4.13}$$

In view of this, (4.7), (4.8) are equivalent to

$$\sup_{x \in \mathcal{V}_n^{\mathrm{out}}} \left(\inf_{x \in \mathcal{V}_n^{\mathrm{out}}} \right) G^n(x; s) = \frac{n^{1-\gamma}}{s}(1 + o(1)), \tag{4.14}$$

$$\sup_{x \in \mathcal{V}_n^{\mathrm{out}}} G^n(x; s) - \inf_{x \in \mathcal{V}_n^{\mathrm{out}}} G^n(x; s) = o(n^{1-2\gamma}). \tag{4.15}$$

To verify these two claims we need very precise estimates on $q_k^n(\mathbf{0}, x)$. Fortunately, the estimates from [CD02] could be used with a small refinement:

Lemma 4.3 ([CD02] Lemma 3.1). *(a) Let $\varepsilon_n = 1/\sqrt{\log L}$. There is a finite constant C such that*

$$\sup_{k \geq \varepsilon_n L^2} \sup_{x \in \mathcal{V}_n} \varepsilon_n L^2 \rho_k^n(\mathbf{0}, x) \leq C. \tag{4.16}$$

(b) If $s_n \to \infty$ as $n \to \infty$, then

$$\sup_{k \geq s_n L^2} \sup_{x \in \mathcal{V}_n} L^2 |\rho_k^n(\mathbf{0}, x) - L^{-2}| = o(s_n^{-1}). \tag{4.17}$$

(c) If $u_n \to \infty$ as $L \to \infty$, then

$$\sup_{x \in \mathcal{V}_n} \sup_{u_n(1+|x|)^2 \leq k \leq \varepsilon_n L^2} |\pi k \rho_k^n(\mathbf{0}, x) - 1| \to 0. \tag{4.18}$$

(d) There is a finite constant C such that

$$\sup_{k \geq 0, \, x \in \mathcal{V}_n} (1 + |x|^2) \rho_k^n(\mathbf{0}, x) \leq C. \tag{4.19}$$

Proof. The claims (a), (c) and (d) are proved in [CD02]. The claim (b) is a refinement of [CD02], where the same expression was proved to be $o(1)$. As in [Cox89] we use a very precise expansion of $q_k(\mathbf{0}, x)$ from [BRR76], Corollary 22.3, to prove (b). As stated there

$$q_k(\mathbf{0}, x) = \phi_k(x) + \psi_k(x). \tag{4.20}$$

Here $\phi_k(x) = (\pi k)^{-1} \exp(-|x|^2/k)$ with

$$|x| = \min\{\sqrt{(x_1 + z_1 L)^2 + (x_2 + z_2 L)^2} : z \in \mathbb{Z}^2\}, \tag{4.21}$$

and

$$\psi_k(x) = \phi_n(x) \sum_{r=1}^{w} k^{-r/2} B_r(x/\sqrt{k}) + e_w(x, k), \tag{4.22}$$

where w is an integer larger than 2, each $B_r(x)$ is a polynomial of degree at most $3r$, and

$$\sum_{x \in \mathbb{Z}^2} |e_w(x, k)| = o(k^{-w/2}) \qquad (n \to \infty). \tag{4.23}$$

To get (b), it is sufficient to improve estimates (i), (ii) on page 1373 of [CD02], namely to show

$$L^2 \sum_{z \in \mathbb{Z}^2} \phi_{[s_n L^2]}(x + Lz) - L^2 = o(s_n^{-1}), \tag{4.24}$$

$$L^2 \sum_{z \in \mathbb{Z}^2} \psi_{[s_n L^2]}(x + Lz) = o(s_n^{-1}). \tag{4.25}$$

Let $x' = x/L$, then the left-hand side of (4.24) can be written as

$$\sum_{z \in \mathbb{Z}^2} \frac{1}{s_n \pi} \exp\left\{-\frac{|x' + z|^2}{s_n}\right\} - \int_{\mathbb{R}^2} \frac{1}{s_n \pi} \exp\left\{-\frac{|x' + y|^2}{s_n}\right\} dy$$

$$= \frac{1}{\pi} \sum_{z \in \mathbb{Z}^2/\sqrt{s_n}} \int_{I_{s_n^{-1/2}}} e^{-|x'' + z|^2} - e^{-|x'' + z + y|^2} dy, \quad (4.26)$$

where $x'' = x'/\sqrt{s_n}$ and $I_\delta = [-\delta/2, \delta/2]^2$. Let $h(u) = e^{-|u|^2}$. It is easy to show using Taylor expansion up to forth order and eliminating the odd terms that

$$\int_{I_\delta} h(u + v) - h(v) \, dv = \int_{I_\delta} \left[\frac{1}{2}\left(\partial_{11} h(u) v_1^2 + \partial_{22} h(u) v_2^2\right) + R(u)\left(\frac{\delta}{2}\right)^4\right] dv,$$

$$= \frac{\delta^4}{8}(\partial_{11} h(u) + \partial_{22} h(u)) + R(u)\delta^6.dv \tag{4.27}$$

where the reminder $R(u)$ is bounded by the sum of suprema of all derivatives of forth order over the set $u + I_\delta$. Since

$$\sum_{z \in \mathbb{Z}^2/\sqrt{s_n}} \partial_{11} h(z) + \partial_{22} h(z) + R(z) < \infty \tag{4.28}$$

we get that (4.26) is bounded from above by $C(\sqrt{s_n})^{-4} = C s_n^{-2}$ uniformly in $x \in \mathcal{V}_n$. This proves (4.24).

To show (4.25) is much simpler. It follows from the calculation at page 1343 of [Cox89] that the contribution of the first part of ψ (see (4.22)) is actually bounded by

$CL^{-1} \ll s_n^{-1}$. Similarly, (4.23) yields that the contribution of the second part of ψ is bounded by

$$L^2 \sum_{z \in \mathbb{Z}^2} |e_w(x + zL, s_n L^2)| \le L^2 \sum_{z \in \mathbb{Z}^2} |e_w(z, s_n L^2)| = L^2 o(L^{-2} s_n^{-w/2}). \tag{4.29}$$

It is now sufficient to take $w = 3$ to finish the proof of (4.25) and therefore of Lemma 4.3. $\qquad\square$

We can now finish the proof of (4.14) and (4.15). Choose s_n such that $n^\gamma \ll s_n \ll n^{1/2-2\gamma} \ll n^{1-\gamma}$, which is possible for all $\gamma \in (0, 1/6)$. Then, first, by Lemma 4.3(b), uniformly in x,

$$\sum_{k=s_n L^2}^{\infty} e^{-sk/h} \rho_k^L(\mathbf{0}, x) = \sum_{k=s_n L^2}^{\infty} e^{-sk/h} L^{-2} (1 + o(s_n^{-1}))$$
$$= L^{-2} e^{-ss_n/n^{1-\gamma}} (1 - e^{-s/h})^{-1} (1 + o(s_n^{-1})) \tag{4.30}$$
$$= \frac{n^{1-\gamma}}{s} (1 + o(n^{-\gamma})).$$

Second, by Lemma 4.3(a), uniformly in $x \in \mathcal{V}_n^{\mathrm{out}}$ (recall $h = 2^{2n} n^{1-\gamma}$),

$$\sum_{k=\varepsilon_n L^2}^{s_n L^2} e^{-sk/h} \rho_k^L(\mathbf{0}, x) \le C e^{-sk/h} \frac{C}{\varepsilon_L L^2} \le \frac{C s_n}{\varepsilon_n s} = o(n^{1-2\gamma}). \tag{4.31}$$

Third, by Lemma 4.3(c),

$$\sum_{k=u_n(1+|x|)^2}^{\varepsilon_n L^2} e^{-sk/h} \rho_k^L(\mathbf{0}, x) \le \sum_{k=u_n(1+|x|)^2}^{\varepsilon_n L^2} \frac{C}{k} \le C\left[\log \frac{\varepsilon_n}{u_n} - 2\log \frac{1+|x|}{L}\right]. \tag{4.32}$$

Since $1 \ge (1 + |x|)/L > n^{-\kappa}$ for $x \in \mathcal{V}_n^{\mathrm{out}}$, choosing $u_n = n^{1-3\gamma}$, the last expression is bounded by $C' \log n \ll n^{1-2\gamma}$. Finally,

$$\sum_{k=0}^{u_n(1+|x|)^2} e^{-sk/h} \rho_k^L(\mathbf{0}, x) \le \sum_{k=0}^{u_n(1+|x|)^2} C(1 + |x|^2)^{-1} \le C u_n \ll n^{1-2\gamma}. \tag{4.33}$$

Putting together (4.30)–(4.33) we get that for all $x \in \mathcal{V}_n^{\mathrm{out}}$

$$G^n(x; s) = \frac{n^{1-\gamma}}{s} (1 + \bar{o}(1)) + o(n^{1-2\gamma}), \tag{4.34}$$

which is equivalent to (4.14), (4.15). (Here again $\bar{o}$ denotes an error that is independent of $x \in \mathcal{V}_n^{\mathrm{out}}$.) This finishes the proof of Proposition 4.2. $\qquad\square$

As in the case of the REM, the assumptions of Proposition 4.2 are verified for a Poisson cloud on the torus:

Lemma 4.4. *Let A_n be Poisson clouds with densities ρ_n satisfying $\lim_{n\to\infty} \rho_n 2^{2n} n^{-\gamma} = \rho \in (0, \infty)$ for some $\gamma \in (0, 1)$. Then*
   (i) $\mathbb{P}$-a.s. $n^{-\gamma}|A_n| \in (\rho - \delta, \rho + \delta)$ for any $\delta > 0$ and n large enough.
   (ii) Let $\gamma \in (0, 1)$. Then for any $\kappa > 1 + \gamma$, $\mathbb{P}$-a.s. for n large enough

$$d_{\min} := \min\{d(x, y) : x, y \in A_n\} \ge 2^n n^{-\kappa}. \tag{4.35}$$

(iii) If $\gamma \in (0, 1/6)$, then the set A_n satisfies $\mathbb{P}$-a.s. the assumptions of Proposition 4.2 for all n large enough.

(iv) The claims *(ii)*, *(iii)* stay valid if the set A_n is replaced by $A_n \cup \{0\}$.

Proof. The proof follows the same lines as the proof of Proposition 3.2 and is left to the reader. $\square$

Corollary 4.5. *If $\gamma \in (0, 1/6)$ then Condition (C) holds for $r(n)$, $\rho(n)$ from (4.3).*

4.2. **Proof of Theorem 4.1.** It remains to verify Conditions (B), (D), 5 and 6 for objects in (4.3). The condition (A) follows trivially from (4.1).

Condition (D). We need first an estimate on the Green's function $G_{\xi_n}^n(0, x)$. Since (up to a very small error due to smoothing (4.9) that we can ignore) $G_{\xi_n}^n(0, x) = \sum_{k=0}^{\xi_n} q_k^n(0, x)$, we use Lemma 4.3. First, by (b) of this lemma,

$$\sum_{k=s_n L^2}^{\xi_n} q_k^n(0, x) \le \sum_{k=s_n L^2}^{\xi_n} CL^{-2} \le C\xi_n L^{-2} \le Cmn^{1-\gamma}. \tag{4.36}$$

Second, by (a),

$$\sum_{k=\varepsilon_n L^2}^{s_n L^2} q_k^n(0, x) \le (s_n - \varepsilon_n)L^2 \frac{C}{\varepsilon_n L^2} \le Cs_n \varepsilon_n^{-1} \ll n^{1-\gamma} \tag{4.37}$$

if s_n is such that $s_n \to \infty$ and $s_n \ll n^{1/2-\gamma}$. Further, by (c),

$$\sum_{k=u_n(1+|x|)^2}^{\varepsilon_n L^2} q_k^n(0, x) \le \sum_{k=u_n(1+|x|)^2}^{\varepsilon_n L^2} \frac{C}{k} \le C \log \frac{L^2 \varepsilon_n}{u_n(1+|x|)^2}, \tag{4.38}$$

and by (d)

$$\sum_{k=0}^{u_n(1+|x|)^2} q_k^n(0, x) \le u_n. \tag{4.39}$$

Taking $u_n = n^{1-\gamma}$ and putting together (4.36)–(4.39) we get that for all $x \in \mathcal{V}_n$

$$G_{\xi_n}^n(0, x) \le C_1 mn^{1-\gamma} + C_2 \log \frac{L^2 \varepsilon_n}{u_n(1+|x|)^2}. \tag{4.40}$$

We can now prove Condition (D) for $\lambda_n = n^{\delta-1}$ with $0 < \delta < \gamma$. We first treat $x \in \mathcal{V}_n$ with $d(0, x) \ge 2^n n^{-\kappa}$, $\kappa > 1$. In this case $G_{\xi_n}^n(0, x) \le C_1 mn^{1-\gamma}$ and thus $\lambda_n G_{\xi_n}^n(0, x) \ll 1$. Therefore

$$\sum_{x \in \mathcal{V}_n : |x| \ge 2^n n^{-\kappa}} \left(e^{\lambda_n G_{\xi(n)}^n(0,x)} - 1 \right) \le \sum_{x \in \mathcal{V}_n : |x| \ge 2^n n^{-\kappa}} 2\lambda_n G_{\xi(n)}^n(0, x) \le 2\lambda_n m r(n). \tag{4.41}$$

The rest of the sum is negligible. Indeed,

$$\sum_{x\in\mathcal{V}_n:|x|\leq 2^n n^{-\kappa}} e^{\lambda_n G^n_{\xi(n)}(0,x)} \leq e^{C_1 mn^{\delta-\gamma}} \sum_{|x|\leq 2^n n^{-\kappa}} \left(\frac{c2^{2n}n^{1/2-\gamma}}{(1+|x|^2)}\right)^{C_2\lambda_n}$$

$$\leq (c2^{2n}n^{1/2-\gamma})^{C_2\lambda_n}\int_1^{2^n n^{-\kappa}} z^{1-2C_2\lambda_n}\,\mathrm{d}z \tag{4.42}$$

$$\leq C(c2^{2n}n^{1/2-\gamma})^{C_2\lambda_n}(2^n n^{-\kappa})^{2-2C_2\lambda_n} \leq C2^{2n}n^{-2\kappa} \ll \lambda_n r(n).$$

Condition (B). This condition follows from the next lemma.

Lemma 4.6. *Uniformly for* $x \in T_\varepsilon^M$

$$\lim_{n\to\infty} n^{-1}G^n_{T_\varepsilon^M\setminus\{x\}}(x,x) = \frac{2\log 2}{\pi}. \tag{4.43}$$

Proof. By Lemma 4.4 there is no point of the top in the disk $D_x(2^n n^{-\kappa})$ with radius $2^n n^{-\kappa}$ around $x \in T_\varepsilon^M$. Therefore, by e.g. Theorem 1.6.6 of [Law91],

$$G^n_{T_\varepsilon^M\setminus\{x\}}(x,x) \geq G^n_{D_x(2^n n^{-\kappa})^c}(x,x) = \frac{2}{\pi}\log(2^n n^{-\kappa}) + O(1) = \frac{2\log 2}{\pi}n + o(n). \tag{4.44}$$

Further, after hitting the boundary of $D_x(2^n n^{-\kappa})$ the simple random walk Y_n has probability of order $n^{-\gamma}$ to return back to x before hitting $T_\varepsilon^M \setminus \{x\}$. Therefore

$$G^n_{T_\varepsilon^M\setminus\{x\}}(x,x) \leq \sum_{k=0}^\infty n^{-\gamma k}G^n_{D_x(2^n n^{-\kappa})^c}(x,x) \leq \frac{2\log 2}{\pi}n + o(n). \tag{4.45}$$

This finishes the proof. $\qquad\square$

Condition 6. This condition can be easily verified as in the REM case using Lemma 4.4 and Proposition 4.2.

Condition 5. The proof of the last condition is slightly more complicated, it follows the lines of the proof of Lemma 7.4 in [BČM06]. For any $\delta > 0$ we have defined event $A_n(\delta)$ by (see (2.13))

$$\exists 1 < j_n < \zeta_n : S_n(r_n(j_n)) \leq t'_n \leq S_n(r_n(j_n+1)) - \delta t(n), \tag{4.46}$$

where $t(n)/2 \leq t'_n \leq (1+\theta)t(n)$ is a deterministic time sequence.

We want to show that $\mathbb{P}[X_n(t'_n) = U_n(j_n)|A_n(\delta),\boldsymbol{\tau}] > 1 - \delta$ for all n large enough. To simplify the notation we define $y = U_n(j_n)$, $u_n = S_n(r_n(j_n))$. We further fix $\omega = (\gamma+7)/(1-\alpha)$ (this value has no particular importance, any larger value would work), and $\lambda = \gamma + \omega\alpha/2 + 1$.

By Condition 1 it is possible to choose ε small enough such that the mean time spent in T^ε before ξ_n is smaller than $\delta^3 t(n)$. This implies (using Chebyshev inequality) that

$$\mathbb{P}\big[\exists v_n \in [S_n(r_n(j_n+1)) - \delta t(n), S_n(r_n(j_n+1))] :$$

$$X_n(s_n) = y|A_n(\delta),\boldsymbol{\tau}\big] \geq 1 - c\delta^2. \tag{4.47}$$

It is easy to show that conditionally on the event in the last display, the probability that the process X_n leaves disk $\mathcal{D} := D_y(2^n n^{-\lambda})$ between times u_n and v_n tends to 0 as n increases. Therefore,

$$\mathbb{P}\big[X_n(t') \in \mathcal{D} \,\forall t' \in [u_n, v_n] | A_n(\delta), \boldsymbol{\tau}\big] \geq 1 - c\delta^2. \tag{4.48}$$

We use B_n to denote the event in the last display. To finish the proof of Condition 5 it is sufficient to show that for all n large

$$\mathbb{P}[X_n(t'_n) = y | \boldsymbol{\tau}, B_n, u_n, y] \geq 1 - \delta/2. \tag{4.49}$$

The Markov process $(X_n(u_n + s) : s \in [0, v_n - u_n])$, given $\boldsymbol{\tau}$, B_n, u_n, y is equal in law to the process $(V(s) : s \in [0, v_n - u_n])$ conditioned on the event $\{T > v_n - u_n\}$ where V and T are constructed as follows:

(i) V stays at site y for an exponential, mean τ_y, amount of time, then

(ii) with probability $p(n)$, the probability that a random walk starting at y escapes $\mathcal{D}$ before returning to site y, the process terminates and T is the termination time. With probability $1 - p(n)$ the process V performs an excursion away from y conditioned not to leave $\mathcal{D}$. At the end of the excursion it returns to y and step (i) resumes and so on.

The important point is that the number $p(n)$ is of order $1/n$ while (recall $y \in T_\varepsilon^M$) the mean time spent at y per visit exceeds $\varepsilon 2^{2n/\alpha} n^{-\gamma/\alpha}$. Thus the conditioning event has probability bounded below by $C(\varepsilon, \theta)$. Hence it will suffice to show that $\mathbb{P}[\bar{V}(t'_n - u_n) \neq y | \boldsymbol{\tau}]$ tends to zero as n tends to infinity $\boldsymbol{\tau}$-a.s. where process $(\bar{V}(u) : u \geq 0)$ is a Markov process that alternates staying at site y an exponential amount of time with mean τ_y and performing excursions away from y conditioned to stay within $\mathcal{D}$ (again staying at each site a time according to $\boldsymbol{\tau}$).

We first show that $\boldsymbol{\tau}$-a.s. for n sufficiently large, the expected duration of a conditioned excursion from y is very small compared to τ_y uniformly over possible $y \in T_\varepsilon^M$. It is easy to prove that in $\mathcal{D}$ there are only traps shallower than $\varepsilon n^{-\omega} 2^{2n/\alpha} n^{-\gamma/\alpha}$. Indeed, let $\mathcal{B}(y)$ be the event

$$\mathcal{B}(y) = \left\{ y \in T_\varepsilon^M, \exists x \in \mathcal{D}, \tau_x \geq \varepsilon n^{-\omega} 2^{2n/\alpha} n^{-\gamma/\alpha} \right\}. \tag{4.50}$$

Then,

$$\mathbb{P}[\bigcup_{y \in \mathcal{V}_n} \mathcal{B}(y)] \leq C 2^{2n} (2^{-2n} n^\gamma)^2 n^{\omega\alpha} 2^{2n} n^{-2\lambda} = C n^{-2} \tag{4.51}$$

and the claim follows easily by the Borel-Cantelli lemma.

It was proved, e.g. in [BČM06], that the expected number of visits to $z \in \mathcal{D} \setminus \{y\}$ during an excursion that does not leave the disk is smaller than 2. The expected duration of the i-th excursion, V_i, thus satisfies

$$\mathbb{E}[V_i | \boldsymbol{\tau}] \leq 2 \sum_{z \in \mathcal{D} \setminus \{y\}} \tau_z \leq 2 \sum_{z \in \mathcal{V}_n} \tau_z \mathbb{1}\{\tau_z \leq n^{-\omega} \varepsilon 2^{2n/\alpha} n^{-\gamma/\alpha}\}. \tag{4.52}$$

It follows from (4.1) that

$$\mathbb{E}[\tau_x \mathbb{1}\{\tau(x) \leq u\}] = O(u^{1-\alpha}) \qquad (u \to \infty). \tag{4.53}$$

Therefore, by Chebyshev inequality,

$$\mathbb{P}\left[\sum_{z \in \mathcal{V}_n} \tau_z \mathbb{1}\{\tau_z \leq n^{-\omega} \varepsilon 2^{2n/\alpha} n^{-\gamma/\alpha}\} \geq 2^{2n/\alpha} n^{-\gamma/\alpha} n^{-5} \right] \leq C n^{-2}. \tag{4.54}$$

Hence, for a.e. τ and n large enough

$$\mathbb{E}[V_i|\boldsymbol{\tau}] \leq C 2^{2n/\alpha} n^{-\gamma/\alpha} n^{-5}. \tag{4.55}$$

The expected number of excursions of $\bar{V}$ before time $v_n - u_n = O(t(n))$ is bounded by a multiple of n, the mean of the total time spent by $\bar{V}$ during the interval $[0, v_n - u_n]$ away from y is easily bounded by $C 2^{2n/\alpha} n^{-\gamma/\alpha} n^{-4}$ for C depending on ε but not on n.

We claim that (for n sufficiently large) for any $s \in [0, v_n - u_n]$, $\mathbb{P}[\bar{V}(s) \neq y] \leq 2C/n^2$. Suppose not. Then for some s', $\mathbb{P}[Y(s') \neq y] \geq 2C/n^2$. We have that the expected total time spent by $\bar{V}$ away from y in interval $[s', s' + 2^{2n/\alpha} n^{-\gamma/\alpha} n^{-2}]$ is bounded by $C 2^{2n/\alpha} n^{-\gamma/\alpha} n^{-4}$, so there exists $s'' \in [s', s' + 2^{2n/\alpha} n^{-\gamma/\alpha} n^{-2}]$ so that $\mathbb{P}[\bar{V}(s'') \neq y] \leq C/n^2$. On the other hand, by the Markov property for $\bar{V}$, if ν is the time of the first jump from y,

$$\mathbb{P}[\bar{V}(s'') \neq y] \geq \mathbb{P}[\bar{V}(s'') \neq y \cap \{\nu > s'' - s'\}] > \frac{1}{2}\mathbb{P}[\bar{V}(s') \neq y] \geq C/n^2. \tag{4.56}$$

for n sufficiently large. This contradiction gives the desired result.

REFERENCES

[AC06]   A. Asselah and F. Castell, *Self intersection times for random walks, and random walk in random scenery in dimensions $d \geq 5$*, preprint, 2006.

[BBM05]   G. Ben Arous, L. V. Bogachev, and S. A. Molchanov, *Limit theorems for sums of random exponentials*, Probab. Theory Related Fields **132** (2005), 579–612.

[BBG03a]   G. Ben Arous, A. Bovier, and V. Gayrard, *Glauber dynamics of the random energy model. I. Metastable motion on the extreme states*, Comm. Math. Phys. **235** (2003), no. 3, 379–425.

[BBG03b]   G. Ben Arous, A. Bovier, and V. Gayrard, *Glauber dynamics of the random energy model. II. Aging below the critical temperature*, Comm. Math. Phys. **236** (2003), no. 1, 1–54.

[BČ05]   G. Ben Arous and J. Černý, *Bouchaud's model exhibits two aging regimes in dimension one*, Ann. Appl. Probab. **15** (2005), no. 2, 1161–1192.

[BČ06]   G. Ben Arous and J. Černý, *Dynamics of trap models*, to appear in Les Houches summer school, no. LXXXIII, 2006.

[BČM06]   G. Ben Arous, J. Černý, and T. Mountford, *Aging for Bouchaud's model in dimension two*, Probab. Theory Related Fields **134** (2006), no. 1, 1–43.

[BG06]   G. Ben Arous and V. Gayrard, *Elementary potential theory on the hypercube*, preprint, 2006.

[Ber96]   J. Bertoin, *Lévy processes*, Cambridge Tracts in Mathematics, vol. 121, Cambridge University Press, Cambridge, 1996.

[BRR76]   R. N. Bhattacharya and R. Ranga Rao, *Normal approximation and asymptotic expansions*, John Wiley & Sons, New York-London-Sydney, 1976, Wiley Series in Probability and Mathematical Statistics.

[Bou92]   J.-P. Bouchaud, *Weak ergodicity breaking and aging in disordered systems*, J. Phys. I (France) **2** (1992), 1705–1713.

[BD95]   J.-P. Bouchaud and D. S. Dean, *Aging on Parisi's tree*, J. Phys I(France) **5** (1995), 265.

[Čer03]   J. Černý, *On two properties of strongly disordered systems, aging and critical path analysis*, Ph.D. thesis, EPF Lausanne, 2003.

[Cox89]   J. T. Cox, *Coalescing random walks and voter model consensus times on the torus in $\mathbf{Z}^d$*, Ann. Probab. **17** (1989), no. 4, 1333–1366.

[CD02]   J. T. Cox and R. Durrett, *The stepping stone model: new formulas expose old myths*, Ann. Appl. Probab. **12** (2002), no. 4, 1348–1377.

[Dia88]   P. Diaconis, *Group representations in probability and statistics*, Institute of Mathematical Statistics Lecture Notes—Monograph Series, 11, Institute of Mathematical Statistics, Hayward, CA, 1988.

[FIN02]   L. R. G. Fontes, M. Isopi, and C. M. Newman, *Random walks with strongly inhomogeneous rates and singular diffusions: convergence, localization and aging in one dimension*, Ann. Probab. **30** (2002), no. 2, 579–604.

[GKS06]   N. Gantert, W. König, and Z. Shi, *Annealed deviations for random walk in random scenery*, to appear in Annales de l'Institut Henri Poincaré (2006).

[GvdHK06] N. Gantert, R. van der Hofstad, and W. König, *Deviations of a random walk in a random scenery with stretched exponential tails*, to appear in Stochastic Processes and their Applications (2006).

[KS79]   H. Kesten and F. Spitzer, *A limit theorem related to a new class of self-similar processes*, Z. Wahrsch. Verw. Gebiete **50** (1979), no. 1, 5–25.

[Law91]   G. F. Lawler, *Intersections of random walks*, Probability and its Applications, Birkhäuser Boston Inc., Boston, MA, 1991.

[Mat88]   P. Matthews, *Covering problems for Markov chains*, Ann. Probab. **16** (1988), no. 3, 1215–1228.

[Mat89]   P. Matthews, *Some sample path properties of a random walk on the cube*, J. Theoret. Probab. **2** (1989), no. 1, 129–146.

GÉRARD BEN AROUS, COURANT INSTITUTE OF MATHEMATICAL SCIENCES, NEW YORK UNIVERSITY, 251 MERCER STREET, NEW YORK, N.Y. 10012-1185, AND ÉCOLE POLYTECHNIQUE FÉDERALE DE LAUSANNE, 1015 LAUSANNE, SWITZERLAND

  E-mail address: `gerard.benarous@epfl.ch`

JIŘÍ ČERNÝ, ÉCOLE POLYTECHNIQUE FÉDERALE DE LAUSANNE, 1015 LAUSANNE, SWITZERLAND

  E-mail address: `jiri.cerny@epfl.ch`